6-2018

Dispersion and uncertainty in density forecasts: Evidence from surveys of professional forecasters

You LI
Singapore Management University, you.li.2013@phdecons.smu.edu.sg

DISPERSION AND UNCERTAINTY IN DENSITY FORECASTS:

EVIDENCE FROM SURVEYS OF PROFESSIONAL FORECASTERS

You Li

A DISSERTATION

In

ECONOMICS

Presented to the Singapore Management University in Partial Fulfilment

of the Requirements for the Degree of PhD in Economics

2018

____

Supervisor of Dissertation

____

PhD in Economics, Programme Director

Dispersion and Uncertainty in Density Forecasts:
Evidence from Surveys of Professional Forecasters

by
You Li

Submitted to School of Economics in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Economics

Dissertation Committee:

Anthony Tay (Supervisor/Chair)
Associate Professor of Economics
Singapore Management University

Yiu Kuen Tse
Lee Kong Chian Professor of Economics
Singapore Management University

Hwee Kwan Chow
Professor of Economics and Statistics
Singapore Management University

Aurobindo Ghosh
Assistant Professor of Finance
Singapore Management University

Singapore Management University
2018

Abstract

This dissertation studies patterns of dispersion in density forecasts as reported in surveys of professional forecasters. We pay special attention to the role of uncertainty in explaining dispersion in professional forecasters' density forecasts of real output growth and inflation. We also consider the relationship between survey design and forecaster behavior. The last chapter describes future research exploring the characteristics of forecaster expectations using probability integral transforms.

As a starting point, chapter one gives a summary of the literature that tries to answer, using data from survey of forecasters, the following three key questions: Why do forecasters disagree? What do density forecasts reveal in addition to point forecasts? Does disagreement serve as a good proxy for forecaster uncertainty? This chapter provides an overview of the studies and briefly discusses how this dissertation can make contributions to the literature.

Chapter two explores the role of uncertainty in explaining dispersion in professional forecasters' density forecasts of real output growth and inflation. We consider three separate notions of uncertainty: general macroeconomic uncertainty (the fact that macroeconomic variables are easier to forecast at some time than others), policy uncertainty, and forecaster uncertainty. The main finding is that dispersion in individual density forecasts is related to overall macroeconomic uncertainty and policy uncertainty, while forecaster uncertainty (which we define as the average in the uncertainty expressed by individual forecasters) appears to have little role in forecast dispersion.

Chapter three examines the relationship between survey design and forecasters' behavior by exploiting changes to the probability bins provided to

forecasters at the solicitation of density forecasts. We consider three important surveys, namely Survey of Professional Forecasters by the Philadelphia Fed (US-SPF), Survey of Professional Forecasters by the European Central Bank (ECB-SPF), and Survey of External Forecasters by Bank of England (SEF). While the adjustment of forecast bins can reasonably arise from the fluctuation of underlying macroeconomic variable, there are also cases where the modification is neutral to the economic environment. Our analysis examines how disagreement and forecaster uncertainty respond to these two different categories of survey changes. The results suggest that disagreement only responds to changes caused by real economy. Uncertainty responds to both and the effect is more persistent. These empirical facts highlight the importance of behavioral perspective when inferences are drawn from professional forecasts.

I summarize our conclusion in Chapter four, and describe future research plan exploring the features of forecaster uncertainty using probability integral transforms ("z-statistics"), a commonly-used test for density forecast optimality. We focus on the shape of the distribution of z-statistics, which is informative about the confidence level as well as bias (optimism or pessimism) of forecasters. There is evidence of significant hysteresis and that survey scheme greatly affects the performance of density forecasts.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I want to express my deepest gratitude to Prof. Anthony Tay, who led me to the career in research and provided excellent guidance throughout my Ph.D. years. It would never have been possible for me to take this work to completion without his help and inspiration. My sincere thanks also go to Prof. Yiu Kuen Tse, Prof. Hwee Kwan Chow, Prof. Aurobindo Ghosh and Prof. Yu Sun. They all helped a lot in my research and Ph.D. study.

I thank sincerely my husband and my parents, who supported me throughout these years. I would also want to thank Thor Qiu Ling, Amelia Tan and Yong Fong Koh for their administrative support and my Ph.D. program mates for their kindness and help.

Chapter 1 Introduction and Survey

Expectations play an important role in the management of an economy and more and more central banks are building their own dataset of economic predictions based on surveys of individuals, including professional forecasters. These datasets allow researchers and policy makers to examine the expectation formation process of individuals and thus can help improve economic modelling as well as formulate more robust and effective monetary policies. In addition to point forecasts, many such surveys now collect density forecasts. Density forecasts potentially depict a more complete picture of expectations and thus provide a wider framework to study how individuals form their economic predictions. These well-developed surveys allow both economic analysts and policy makers to better understand the expectation formation process of individuals in the economy. However, there are many features highlighted in the datasets that are quite different from assumptions made in many classical economic models and cultivate a growing literature focusing on the features as well as determinants of these features.

Why do forecasters disagree?

Surveys of macroeconomic forecasts show that forecasters generally disagree with each other. Pervasive evidence of forecast dispersion is found in both point and density forecasts, even among a group of economic agents that should be reasonably homogeneous in terms of ability and motivation. This fact indicates that expectations are actually heterogeneous, contradictory to the widely accepted basic assumption of “rational expectations”, and emphasizes the necessity of taking heterogeneous expectations into account when constructing macroeconomic models for policy and forecasting. Therefore, a large number of studies explore the determents of disagreement and explanations considering different types of heterogeneity among forecasters have been proposed to answer this question. The first one is heterogeneity in terms of information set. This relates to information rigidity literature including noisy information model (Sims, 2002) where the dispersion stems from uncertain news, and sticky information model proposed by Mankiw and Reis (2002), where forecasters only occasionally update their information at different rate due to the cost in updating news. Coibion and Gorodnichenko (2012) provide empirical evidence to support the sticky information model. Another form of heterogeneity is linked to the models used by the forecasters. Branch (2004,2007) finds that forecasters dynamically switch between a set of costly alternative prediction models in each period. In addition to model uncertainty, Patton and Timmermann (2010) and Lahiri and Sheng (2010) show that difference in priors is an important determinant of dispersion, especially at longer horizons. Manzan (2011) extends the literature by illustrating that heterogeneity in interpretation of news as well could explain the source of dispersion. Last but not least, Capistran and Timmermann (2009) examine the role of loss function and find that heterogeneity in the asymmetric loss function could lead to dispersion.

Capistran and Ramos-Francia (2010) assess the relationship between dispersion and monetary policy using a panel dataset of multiple countries. They find that the dispersion of long-run inflation expectations is smaller in countries with targeting regimes. A multi-national study of Dovern, Fritsche and Slacalek (2012) find that dispersion of GDP growth is more strongly affected by real factors and dispersion of price reacts to institutional setting of monetary policy, in particular, central bank independence. These findings indicate that the degree of dispersion may be affected by monetary policy and real economic factors.

What do density forecasts reveal in addition to point forecasts?

Density forecasts are nowadays commonly included in professional forecaster surveys to depict a more complete picture of the forecast uncertainty. Studies of density forecasts have up to now largely focused on whether density forecasts are optimal. One exception is Engelberg, Manski and Williams (2009), who examines whether the central location of density forecasts matches the point forecasts.

One difficulty that economic researchers face is the inability to measure forecaster uncertainty. Since density forecasts appear to provide a direct self-represented measure of forecaster uncertainty, a new and rapidly growing strand of literature studies the second moment of density forecasts and focuses on the features of forecaster uncertainty. Boero, Smith and Wallis (2015) find that there is great heterogeneity and persistence in forecaster uncertainty.

Does disagreement serve as a good proxy for uncertainty?

Declining dispersion from long to short horizons suggests a link between dispersion and uncertainty. Dispersion has been widely taken as a substitute for uncertainty, which has always been the focus of policy makers and difficult to quantitatively measure. However, this approach is subject to a long-lasting debate in the literature. On one side, Zarnowitz and Lambros (1987) and Giordani and Soderlind (2003) find that the dispersion and uncertainty are highly correlated so that dispersion can serve as a reasonable proxy for uncertainty. On the other side, Rich and Tracy (2008) argue that there is little evidence supporting this relationship. Furthermore, Boero, Smith and Wallis (2008, 2015) examine this topic using the Bank of England Survey of External Forecasters and argue that dispersion is not a good proxy for uncertainty, as least during stable periods. Lahiri and Sheng (2010) give opposite argument that dispersion is a reliable measure for uncertainty in a stable period but less useful as a proxy in periods with large volatility.

In this dissertation, I study these topics using three world-wide surveys, namely Survey of Professional Forecasters by the Philadelphia Fed (US-SPF), Survey of Professional Forecasters by the European Central Bank (ECB-SPF), and Survey of External Forecasters by Bank of England (SEF), to further explore the features of density forecasts and underlying determinants of the features.

Chapter two explores the role of uncertainty in explaining dispersion in professional forecasters' density forecasts of real output growth and inflation. We consider three separate notions of uncertainty: general macroeconomic uncertainty (the fact that macroeconomic variables are easier to forecast at some time than others), policy uncertainty, and forecaster uncertainty. We find that dispersion in individual density forecasts is related to overall macroeconomic uncertainty and policy uncertainty, while forecaster uncertainty (which we define as the average in the uncertainty expressed by individual forecasters) appears to have little role in explaining forecast dispersion. On one hand, our work contributes to the literature by extending the analysis on dispersion from point forecasts to density forecasts. Though there is growing literature focusing on density forecast evaluation and forecaster uncertainty, very few studies examine the dispersion pattern in density forecasts. On the other hand, we exploit various forms of uncertainty to explain forecast dispersion, including both subjective and objective measures showing different perspective of the economic environment. By doing so, we are able to add new elements to the literature examining determinants of dispersion. Furthermore, further analysis on forecaster uncertainty and dispersion in this study highlights different features of these two important measures represented in density forecasts and helps us better understand the relationship of them.

Chapter three examines the relationship between survey design and forecasters' behavior by exploiting changes to the probability bins provided to forecasters at the solicitation of density forecasts. While the adjustment of forecast bins can reasonably arise from the fluctuation of underlying macroeconomic variable, there are also cases where the modification is neutral to the economic environment. Our analysis examines how disagreement and forecaster uncertainty respond to the two different categories of survey changes. The results suggest that disagreement only responds to changes caused by real economy. Uncertainty responds to both and the effect is more persistent. This chapter contributes to the literature on characteristics of forecaster uncertainty from the behavioral perspective. More importantly, it provides a note of caution on the survey design: as the survey design itself could affect the forecasters' behavior, we should pay special attention on the form of the survey questions and be very cautious when making any modifications.

Chapter four concludes, and describes a useful avenue for future research. We propose to further investigate the features of forecaster uncertainty using probability integral transforms ("z-statistics"), a commonly-used test for density forecast optimality. We focus on the shape of the distribution of z-statistics, which is informative about the confidence level as well as bias (optimism or pessimism) of forecasters. There is evidence of significant hysteresis and that survey scheme greatly affects the performance of density forecasts. Such a study would provide a new application of the traditional density forecast evaluation method and add to the literature an innovative prospective on analyzing the determinants of key features of density forecasts.

Chapter 2 The Role of Macroeconomic, Policy, and Forecaster Uncertainty in Forecast Dispersion[1]

2.1 Introduction

Surveys of macroeconomic forecasts show that forecasters generally disagree with each other. Point forecasts are typically dispersed, with patterns like those shown in the left panel of Figure 2.1. This panel displays point forecasts for year 1996 PGDP inflation made by forecasters surveyed by the Philadelphia Fed in their quarterly Survey of Professional Forecasters (SPF) over the period 1995Q1 to 1996Q4, i.e., at horizons 7 down to 0 quarters. The point forecasts are dispersed at long horizons, with dispersion falling as the forecaster approaches full realization of the forecast event. Some persistence in the forecasts can also be observed, with relatively optimistic and pessimistic forecasts tending to remain so. Similar patterns have been observed in many other similar datasets, including surveys carried out by Consensus Economics and the European Central Bank (ECB). Explanations put forward for these and other patterns in dispersion include the use of different information sets by forecasters, perhaps due to different degrees of information rigidities among them (Mankiw, Reis and Wolfers 2003), different interpretation of information by forecasters (Kandel and Zilberfarb 1999; Manzan 2011), different loss functions among forecasters (Capistran and Timmermann 2009), and different priors held by forecasters regarding the unconditional distribution of the variables being forecasted (Patton and Timmermann 2010).

---

[1] Coauthored with Anthony Tay.

Several surveys of macroeconomic forecasts also elicit density forecasts and these too tend to be dispersed, as can be seen from the figures in the right column of Figure 2.1. These figures show the median, central 90% interval, and a skewness measure of individual density forecasts for the same variable as the point forecast in the left panel. Density forecasts, of course, offer us the potential of observing a forecaster's expectations in a more complete form. The spread of a density forecast would seem a reasonable direct measure of the level of uncertainty perceived by the forecaster. Asymmetries in the density forecasts, for a given level of uncertainty, may indicate some degree of optimism or pessimism. Dispersion patterns in the SPF density forecasts have been studied by Lahiri and Liu (2006) who explore in particular the determinants of inflation forecast uncertainty. Dispersion patterns in the density forecasts elicited via the Bank of England's Survey of External Forecasters have been documented in Boero, Smith, and Wallis (2008, 2015).

Pervasive evidence of (point and density) forecast dispersion, even among a group of economic agents who should be reasonably homogeneous in terms of ability and motivation, suggest that expectations may in general be heterogeneous. If this is so, then heterogeneous expectations should be taken into account when constructing macroeconomic models for policy and forecasting. But why are expectations dispersed, and does uncertainty play any role? Declining dispersion from long to short horizons suggests a link between dispersion and uncertainty, though the nature of this link is not at all clear. According to forecasting theory, this pattern is somewhat perplexing. At long horizons, point forecasts should be close to the unconditional mean if forecasters have a mean squared error loss function. Dispersion at longer horizons could therefore imply that different forecasters have various loss functions (Capistran and Timmermann 2009). Or it might be simply that different forecasters have different priors regarding the unconditional distribution of the variables being forecast (Patton and Timmermann 2010). If information is interpreted differently or absorbed at different rates, then we might expect greater dispersion at shorter horizons.

In this paper, we explore the role of uncertainty in explaining forecast dispersion, making a distinction between forecaster uncertainty and general uncertainty in the macroeconomic environment. While we might expect the two to be related, they are not necessarily identical. It is generally accepted that there is time-variation in the volatility of macroeconomic variables, so that these variables are easier to predict in some periods, and harder to do so during others. However, it is not difficult to imagine an overconfident forecaster who always issues very narrow density forecasts. As part of the broader concept of uncertainty in the macroeconomic environment, one might also include policy uncertainty, which again might be expected to be related to, but not identical to forecaster uncertainty or variations in predictability. We use density forecasts to construct a measure of forecaster uncertainty, and take advantage of recently developed indices of macroeconomic uncertainty, which focus on whether the economy has become more or less predictable (Jurado, Ludvigson and Ng 2015), and of policy uncertainty, that rely on the prevalence of 'uncertainty' keywords in news articles (Baker, Bloom, and Davis 2016). We correlate our measures of dispersion in density forecasts with these direct measures of uncertainty.

Our work differs from much of the forecast dispersion literature in that we explore dispersion of density forecasts, and not point forecasts. While there are several papers that have documented dispersion of density forecasts (Boero, Smith and Wallis, 2008), the literature has in general focused on explaining the behavior of individual uncertainty (Lahiri and Liu 2006). Our interest is in characterizing the behavior of dispersion in the location, spread, and skewness of individual density forecasts, using the overall levels of these as well as indices of macroeconomic and policy uncertainty as explanatory variables. Whereas most studies focus on inflation expectations, we also study output growth expectations; it turns out that there are some interesting differences in the behavior of the two. This paper is also closely related to research that is aimed at establishing whether or not point forecast dispersion is a good proxy for forecaster uncertainty (Zarnowitz and Lambros, 1987; Giordani and Soderlind, 2003; Rich and Tracy, 2010; Boero, Smith and Wallis, 2008, 2015), which boils down to asking if dispersion can explain individual uncertainty (the general consensus appears to be, mostly, 'no'.) The objective of our exercise, on the other hand, is to see if various forms of uncertainty can explain forecast dispersion, which we take to be a reflection of heterogeneity in expectations. We find, in general, that dispersion is correlated to macroeconomic uncertainty, less so with policy uncertainty, and not correlated with forecaster uncertainty.

In the next section, we describe the SPF survey dataset briefly, focusing on the density forecasts and the percentile-based summary statistics that we use to characterize them. We also describe the patterns of dispersion in these summary statistics. We take a closer look at forecaster uncertainty in Section 2.3, and its relationship to the macroeconomic uncertainty and policy uncertainty indices. Our main results regarding dispersion are reported in Section 2.4, and Section 2.5 concludes.

2.2 Characteristics and Dispersion of Density Forecasts

2.2.1 Data

Our expectations data are forecasts elicited from professional forecasters by the Philadelphia Fed through their Survey of Professional Forecasters. Every quarter the Philadelphia Fed surveys a panel of professional forecasters for their expectations regarding a range of macroeconomic variables at various forecast horizons. The survey is sent out after the release of advanced estimates for the variables for the previous quarter. The variables for which point forecasts are collected include quarterly and annual frequency real and nominal GDP, unemployment, 3-month treasury bill and 10-year treasury bond rates, price indices (GDP price index, CPI and PCE indices), among others. Besides point forecasts, the surveys also elicit density forecasts for growth in the annual averages of real GDP, the GDP price index (PGDP), core CPI and core PCE, and the civilian unemployment rate. The density forecasts take the form of histograms; forecasters are given a set of intervals and asked to provide for each interval an estimate of the probability with which the variable's realization is expected to appear in that interval. Figure 2.2 displays an individual forecaster's density forecasts from the 2009Q1 and 2010Q1 surveys, for growth in the annual-average of real GDP for the year 2010.

Our focus in this paper is on the density forecasts for the annual average of real GDP growth and PGDP inflation, since density forecasts for the other variables are recent additions to the survey (2007Q1 for CPI and PCE inflation, 2009Q2 for unemployment). Density forecasts are elicited for the annual outcomes for the current year and the following year, so for each target year we have forecasts made at horizons of $h = 0$ to $h = 7$ quarters. The Q1 surveys contain forecasts 3- and 7-quarters ahead (corresponding to the current year and following year forecasts respectively), the Q2 surveys contain forecasts at 2- and 6-quarter horizons, the Q3 survey contains forecasts 1- and 5-quarters ahead, and the Q4 survey contains forecasts at 0- and 4-quarter horizons. From 2009Q2 onwards, the survey began requesting, for certain variables, density forecasts for the current and next three years, so in more recent surveys we have forecasts up to 15 quarters ahead. However, for this study we only consider forecasts made 0 to 7 quarters ahead.

We restrict our analysis to forecasts starting with the 1992Q1 survey, even though point and density forecasts for output and inflation are available all the way back to the first survey in 1968Q4. There are several reasons for using the post-1992Q1 sample period. First, there were several definition changes to the variables forecasted prior to the 1992Q1 survey (for output, from nominal GNP to real GNP to real GDP). In some years prior to 1992Q1 the survey asked forecasts for the previous and current year, instead of the current and following year. Since 1992Q1 the definitions have been stable. Second, in the surveys in the 1980s, the interval widths given for density forecasts were switched from one percentage point intervals to two percentage point intervals, leading to much cruder density forecasts. In addition to this, the intervals provided in some of the earlier surveys were sometimes completely misaligned with the expectations of the forecasters, resulting in density forecasts with probabilities concentrated in the first or last open-ended bins. In contrast, the sample period that we use is much cleaner in terms of variable and forecast definitions, and have fewer instances of 'misaligned' bins, and only very few instances where the percentile-based summary statistics that we use cannot be computed. Our sample period ends with the 2016Q4 survey.

Finally, we limit our sample to include only density forecasts where the median of a density forecast matches the forecaster's point forecast. As noted earlier, one possible reason for disagreement among forecasters is simply that they have different loss functions. By matching point and density forecasts, we control for major differences in loss function by limiting ourselves to forecasters with symmetric loss functions, where the point forecasts should (more or less) coincide with the mean or median of the density forecast. The matching method we use is based on the bounds implied by the density forecasts (Engelberg, Manski and Williams 2009). The first step to constructing the matched sample is to calculate the lower and upper bounds of both subjective median and mean. The interval in which the median lies can be obtained from the probabilistic responses directly. To calculate lower and upper bounds on the subjective mean, we assume that each bin's probability mass is placed at the bin's lower and upper endpoint respectively. The results are then generated by averaging the lower and upper endpoints weighted by the probabilities. If the point forecast is located within any of the two sets of bounds, the density forecast is counted as 'matched' to the point forecast. The final matched sample includes 5784 observations for PGDP and 6260 observations for RGDP which means that we retain roughly 30 forecasters in each quarter. The matching takes care of another issue in the sample, that is the presence of outliers and unusual observations that appear to be errors of some sort, or at least difficult to otherwise justify. These outliers are removed via the matching process. Figure 2.3 shows the scatter plots of the median of density forecasts and point forecasts in both the full and the matched samples.

2.2.2 Descriptive Statistics for Density Forecasts

We summarize each density forecast using its median, central 90% range, and a Bowley-type skewness statistic to describe the location, spread, and shape features of the density forecast respectively. The Bowley statistic that we use to measure skewness in the density forecasts is

$$S = \frac{(x_{95} - x_{50}) - (x_{50} - x_5)}{x_{95} - x_5}$$

where x_α represents the α-th percentile of the density forecast. Bowley skewness statistics are usually calculated using the median and the interquartile range, but here we use instead the median x_{50}, and the central 90% range $x_{95} - x_5$. The interquartile version of the Bowley statistic is usually applied to a sample of observations, where the 5th and 95th percentiles are often not meaningful unless the sample size is large. In our application, we use the Bowley statistic to describe a density forecast rather than a sample of data, and using the central 90% range is feasible, and preferred, as it covers more of the distribution. Whereas the range might be considered a measure of individual uncertainty, the Bowley statistic might be interpreted as a direct measure of optimism/pessimism of the forecaster.

Our main reason for using percentile-based descriptions of the density forecasts is that the end bins are open-ended, which complicate the computation of moment-based and entropy-based statistics. Using percentile-based descriptions avoids the strong assumptions which are needed for moment- and entropy-based measures. Of course, the percentile-based statistics that we use also has its disadvantages, e.g., it requires interpolation within the bins (we use linear interpolation of the cumulative probabilities, thus assume probabilities to be evenly spread within each bin). Furthermore, the 5th (95th) percentiles cannot be computed if the probabilities reported for the first or last bins are greater than 5 (probabilities are reported out of 100). While our assumption regarding the shape of the density forecasts is strong, this is mitigated by the fact that the assumption is applied only in the bins in which the 5th, 50th, and 95th percentiles fall. The fact that the 5th and 95th percentiles cannot be computed in some cases is also not a major issue for our sample period. Finally, recent papers have considered moment- and entropy based statistics (Rich and Tracy 2010, Boero, Smith and Wallis 2008, 2015) for some of the issues we examine, so it is interesting to see how our results compare with these studies when using different measures.

We use

$$M_{i,t,h},\ R_{i,t,h}, \text{ and } S_{i,t,h}$$

to denote the median, range, and skewness statistics for forecaster i's period t density forecast of annual GDP growth or annual inflation made h-quarters ahead, $h = 0,1,...,7$. The subscript t is a quarterly date index (1992Q1, 1992Q2, etc.) and represents the survey date. The target year is not represented in this notation, and must be derived from the survey date t and the horizon. We use the sample period $t = 1992q1,...,2016q4$.

We are interested in the dispersion in the three density forecast characteristics among forecasters. Dispersion measures are calculated as standard deviations over the forecasters in each period. We denote our dispersion measures as

$$M^{\sigma}_{t,h} = \text{std.dev.}(M_{i,t,h}),\ R^{\sigma}_{t,h} = \text{std.dev.}(R_{i,t,h}),\ S^{\sigma}_{t,h} = \text{std.dev.}(S_{i,t,h}).$$

We will also be referring to the mean levels of the characteristics, which we denote as

$$M^{m}_{t,h} = \text{mean}(M_{i,t,h}),\ R^{m}_{t,h} = \text{mean}(R_{i,t,h}),\ S^{m}_{t,h} = \text{mean}(S_{i,t,h}).$$

In particular, we treat $R^m_{t,h}$, the average of individual uncertainty, as a measure of overall forecaster uncertainty. We discuss this variable more in the next section, together with measures of macroeconomic and policy uncertainty. The variable $S^m_{t,h}$ is included as an elaboration of average individual uncertainty, indicating asymmetries in relative upside vs downside risks. The variable $M^m_{t,h}$ is included as there may be a relationship between the level of inflation (which should be correlated with the expected level of inflation) and inflation uncertainty (Ball 1992), a relationship which several previous studies have confirmed (Lahiri and Liu 2006, Rich and Tracy 2010). We include $M^m_{t,h}$ for our output growth regressions as well.

2.2.3 Dispersion Patterns in Density Forecasts

Figure 2.4(a) summarizes the behavior of individual PGDP inflation density forecasts, and the dispersion of these forecasts. The top row displays the average of the median, range, and skewness of the individual density forecasts. The bottom row shows the standard deviation of these characteristics of the individual forecasts. Each line in each subfigure corresponds to a target year (1992 to 2017), all plotted against forecast horizon. The top row shows how, on average, the forecasters revise their forecasts each quarter for each target year in our sample, and the bottom row shows the disagreements among the forecasters. Figure 2.4(b) shows the corresponding figures for RGDP growth.

The subfigures marked M^m show moderate revisions to the density forecast medians on average, except for a sharp drop in the RGDP growth forecasts for 2009. The subfigures marked R^m show that average individual uncertainty falls as the forecast horizon approaches zero, with the fall accelerating from horizons 3 to 0 quarters. The accelerating fall should be due in large part to the fact that fewer quarters are being forecasted in these horizons (these are forecasts of annual growth for the year in which the quarterly surveys are taken). We might also expect average uncertainty to fall because more information is (presumably) being incorporated into the forecasts each quarter. This seems more the case for RGDP growth than for PGDP inflation. The subfigure for forecaster uncertainty R^m in PGDP inflation also shows something that might be of concern. There is a systematic drop in average range from the 2014Q1 surveys onwards. This corresponds to a change in bin definitions for PGDP inflation to match those of CPI and core PCE, with the overall range reduced substantially (from "< 0 ",…, "> 8 ", to "< 0 ", …, "> 4"). This is worrying because it might indicate a framing effect as far as the spread of elicited density forecasts are concerned. A much smaller change in the RGDP bin definitions was made in 2009Q2, from "< -2",…, "> 6" to "< -3",…, "> 6". Though it is hard to see from the figures the effects of this change, nonetheless, our regressions will include a new indicator variable $newbin_t$ for both PGDP inflation and RGDP growth. This variable is equal to '0' up to 2013Q4 and '1' thereafter for PGDP inflation forecasts, and '0' up to 2009Q1 and '1' thereafter for RGDP growth forecasts. The subfigures for average skewness S^m show that inflation density forecasts tend to be positively skewed whereas output growth forecasts tend to be negatively skewed. That is, forecasters tend to perceive 'upside risks' for inflation and 'downside risks' for output growth. A major exception is at the end of 2009 when there was a sudden swing to positive skewness in real output growth forecasts. The largest changes in skewness comes in horizons 0 and 1.

The patterns of dispersion in the bottom rows of Figures 2.4(a) and (b) show larger dispersion in the forecast medians at long horizons for both variables, and smaller dispersion at shorter horizons. At the short horizons this might again be due to the fact that there is less to disagree about, but the fall in dispersion seems to be present at all horizons. There appears to be more disagreement at horizon 0 for PGDP inflation than RGDP growth. This has also been noticed in other point forecast datasets (e.g. Patton and Timmermann, 2010), and it is seen here that this regularity extends to density forecasts as well. Dispersion appears to fall slightly for the range, and rise slightly for the skewness, for RGDP growth forecasts as horizon decreases. These patterns are less noticeable for PGDP inflation forecasts. Nonetheless, there is variation over time in the dispersion of all three characteristics of the density forecasts.

2.3 Macroeconomic, Policy, and Forecaster Uncertainty

We have already described our measure of self-reported forecaster uncertainty, namely $R^m_{t,h}$, the average of individual density forecast range $R^m_{i,t,h}$ taken over all forecasters at each survey date. In this section, we discuss recently developed measures of two different notions of uncertainty, namely macroeconomic uncertainty, and policy uncertainty, and explore the relationship between these measures of uncertainty, and forecaster uncertainty. Our main objective, which we will turn to in the next section, is to see how dispersion in density forecasts correlate with these three different "types" of uncertainty.

It is a well-known fact that macroeconomic variables are easier to forecast in some periods than at others because the volatility of their unpredictable components vary over time. Jurado, Ludvigson and Ng (2015) develop an index of macroeconomic uncertainty (which we refer to as MU_t) comprising a weighted average of conditional root mean square forecast errors for a wide range of macroeconomic variables:

$$MU_t = \sum_{j=1}^{N_y} w_j U_{jt}^y(k)$$

$$U_{jt}^y(k) = \sqrt{E[(y_{j,t+k} - E[y_{j,t+k} \mid I_t])^2 \mid I_t]}$$

where k refers to the forecast horizon, and I_t is a large information set on which the forecasts are based. Their set of macroeconomic variables include real output and income, employment, manufacturing and trade sales, consumer spending, housing starts, and many more, totaling 132 variables. The forecast errors for these variables were derived from a factor model utilizing these and 147 financial variables. They calculate macroeconomic uncertainty indices for $k = 1, 3,$ and 12 months. We utilize all three, but report results only for $k = 3$.

Baker, Bloom and Davis (2016) develop an economic policy uncertainty index based on human and automated searches of the archives of ten large newspapers. This index quantifies the volume of relevant news coverage by counting the number of articles related to policy uncertainty starting from January 1985 (monthly average of the standardized number of articles, scaled to an average of 100). We use this data series, downloaded from their website and referred to hereafter as PU_t, as a direct measure of policy uncertainty. We aggregate the monthly index to quarterly frequency by taking the average over each quarter. The left column of Figure 2.5 displays the two indices, where it can be seen that PU_t is considerably more volatile than MU_t. The two series are correlated, but only moderately so, at approximately 0.40, The three uncertainty indices considered in this paper MU_t, PU_t and $R^m_{t,h}$, can be viewed along the objective-subjective spectrum, with MU_t being a purely objective measure, $R^m_{t,h}$ being a purely subjective notion, and PU_t being somewhere in between. We are interested in how dispersion of density forecasts are correlated with these. The right column of Figure 2.5 displays $R^m_{t,h}$, with the higher line representing average forecaster uncertainty relating to forecasts for the following year, and the lower line relating to forecasts for the current year. These figures give a time-series view of $R^m_{t,h}$ whereas the upper middle diagrams in Figures 2.4(a) and (b) show $R^m_{t,h}$ as a function of horizon, for various years. The declining forecaster uncertainty gives rise to the periodicity, especially for the lower horizon forecasts. To explore how subjective forecaster uncertainty relates to the two other measures of uncertainty, we run the regression

$$R^m_{t,h} = \beta_0 + \beta_1 h_{1t} + ... + \beta_7 h_{7t} + \beta_8 newbins_t + \beta_9 R^m_{t-1,h*} + \beta_{10} M^m_{t,h} + \beta_{11} PU_t + \beta_{12} MU_t + \varepsilon_{t,h} \tag{1}$$

for both PGDP inflation and RGDP growth, where $h_{1t},...,h_{7t}$ are horizon dummies. The variable $R^m_{t-1,h*}$, which we will refer to as "lagged individual uncertainty", is included to capture persistence in range across consecutive surveys ($h^* = h+1$ for $h = 0,1,2,4,5,6$ and $h^* = h-3$ for $h = 3,7$). To allow for omitted factors at each survey date, we cluster standard errors by year and quarter, allowing errors to be correlated at horizon pairs (0,4), (1,5), (2,6) and (3,7) within each year. The regression results are displayed in Table 2.1. We also produce 'split-sample' versions of each regression, separating observations associated with 'current-year' forecasts from 'next-year' forecasts.

Overall, there are substantial differences in the results for PGDP inflation and RGDP growth. Some of our results are consistent with previous research using different methodologies. For instance, the coefficients on lagged individual uncertainty show persistence, especially for short-horizon RGDP growth forecasts whereas for PGDP inflation this persistence is weaker and only for longer-horizon forecasts. The coefficients on $M_{t,h}^{m}$ are significant for PGDP inflation (short-horizon). This is consistent with results from previous research which find that positive changes in anticipated inflation is associated with greater uncertainty (Lahiri and Liu 2006, Rich and Tracy 2010).

For PGDP inflation, the correlations between average individual uncertainty and either macro or policy uncertainty seems non-existent, except for a small negative correlation with PU_t for the longer-horizon density forecasts. This suggests little co-movement between forecaster uncertainty and the two other measures of uncertainty. For RGDP growth, the coefficients on PU_t suggest that average forecaster uncertainty is correlated with policy uncertainty for RGDP growth, and that this correlation is mainly at short-horizon forecasts. Average forecaster uncertainty is strongly negatively correlated with MU_t. This is most likely because the largest variation in MU_t in our sample occurred during the 2007-2009 recession, when there was little doubt regarding the state of RGDP growth.

Another interesting set of results can be derived from the horizon dummies. Earlier we noted from the top diagrams of the second columns of Figures 2.4(a) and (b) that range falls with horizon, and that this could be due to less uncertainty because forecasters have more information, or because in horizons 0, 1, and 2, only 1, 2, and 3 quarters worth of growth is being forecasted, whereas for horizons 3 to 7, all four quarters of annual growth are being forecasted. If we assume annual growth to be the sum of four independent quarters of growth $x_1 + x_2 + x_3 + x_4$, each with variance σ^2, then at horizon 0 only x_4 is being forecasted, with variance σ^2, whereas at horizon 1, it is $x_3 + x_4$ that is being forecasted, with variance $2\sigma^2$. If we take the range as equivalent to some multiple of standard deviation, the ratio of the horizon 1 to horizon 0 range should be $\sqrt{2}$. Likewise, the ratio of the horizon 2 to horizon 1 range should be $\sqrt{3/2}$, and the ratio of the horizon 3 to horizon 2 range should be $2/\sqrt{3}$. Thereafter, the ratio of the ranges should be one. The ratios implied by the constant term and horizon dummy coefficients is consistent with this pattern for RGDP growth, though for PGDP inflation forecasts we only observe the fall in range as we move into the lowest horizon.

Finally, the strong significant (negative) coefficient on $newbins_t$ confirms that the reduction in the overall bin range for PGDP inflation that occurred after the 2013Q4 survey led to a large reduction in the range reported by forecasters. This is, of course, obvious from Figure 2.4(a) and Figure 2.6. The smaller (positive) effect in RGDP growth where there was a smaller widening of the overall range provided to the forecasters is also significant at the longer horizons. As mentioned earlier, this may suggest that measures of individual uncertainty derived from density forecasts may be subject to a framing effect. It may of course be the case that both forecast surveyor and forecasters are reacting to the same information. Furthermore, this does not necessarily invalidate our use of direct measures of individual uncertainty from density forecasts, although it does emphasize the need to control for changes in bins offered to the forecasters, and it does warrant caution in the interpretation of self-reported measures of uncertainty, at least as elicited using the methods currently employed in surveys of density forecasts.

We noted earlier that density forecasts for inflation tend to be positive skewed, whereas density forecasts for output growth tend to be negative. We run regressions similar to (2), replacing R^m with S^m, and include R^m as a regressor:

$$\begin{aligned} S^m_{t,h} &= \beta_0 + (\text{horz. dummies}) + \beta_8 newbins_t ... \\ &\quad + \beta_9 S^m_{t-1,h*} + \beta_{10} M^m_{t,h} + \beta_{11} R^m_{t,h} + \beta_{12} PU_t + \beta_{13} MU_t + \varepsilon_{t,h} \end{aligned} \tag{2}$$

The results are reported in Table 2.2. Overall the results are harder to interpret. We leave out the horizon dummies and constant as there is nothing interesting to report. We note a persistence in skewness for both PGDP inflation and RGDP growth density forecasts, and a negative correlation between expected levels and skewness, meaning that as expected levels go up, density forecasts become less skewed (inflation) or more negatively skewed (output growth). Skewness in density forecasts of output growth appears to be negatively correlated with policy uncertainty, whereas skewness in density forecasts of inflation are negatively correlated with macro-uncertainty. Overall, increased uncertainty appears to reduce skewness (inflation) or increase negative skewness (output growth).

2.4 Main Results

We explore in this section the behavior of dispersion in the individual density forecasts, as summarized by their location, range, and skewness statistics $M^\sigma_{t,h}$, $R^\sigma_{t,h}$, $S^\sigma_{t,h}$. We explore in particular the relative degrees to which average forecaster uncertainty ($R^m_{t,h}$) and uncertainty in the overall macroeconomic environment (as measured by MU_t and PU_t) can explain the degree of dispersion observed in these three statistics. We also include the average values of location and skewness, $M^m_{t,h}$ and $S^m_{t,h}$, in the regressions, and lagged dispersion to capture persistence in overall dispersion from one quarter to the next. We include as a further control dispersion in the forecasters' yield spread (nominal rate on 10-year T-bonds minus the nominal rate of 3-month T-bills) in the inflation forecast dispersion regressions, as the year spread is commonly viewed as good predictors of inflation (even if recent evidence suggest that this might not be the case, e.g., Ang, Bekaert, and Wei 2007, Stock and Watson 2009, Rossi and Sekhposyan 2010). The yield spread may also have good predictive power for output growth (Estrella and Hardouvelis 1991, Estrella and Mishkin 1998, Hamilton and Kim 2002), although there is evidence that short-term interest rates forecast output growth better than spreads (Ang, Piazzesi, and Wei 2006). Although we have run the regressions for both the short-rate and the yield spreads for output growth forecast dispersion, we report only the regressions with the short rate, and mention the changes that occur when the spread is used. Finally, dispersions were seen in Figures 2.4(a) and (b) to change systematically with horizon, and we include horizon dummies to allow for this.

$$\begin{aligned} M^{\sigma}_{t,h} &= \beta_0 + (\text{horz. dummies}) + \beta_8 newbins_t + \beta_9 M^{\sigma}_{t-1,h*} \cdots \\ &\quad + \beta_{10} M^m_{t,h} + \beta_{11} R^m_{t,h} + \beta_{12} S^m_{t,h} + \beta_{13} spread^{\sigma} + \beta_{14} PU_t + \beta_{15} MU_t + \varepsilon_{t,h} \end{aligned} \tag{3}$$

The regressions are repeated for $R^{\sigma}_{t,h}$ and $S^{\sigma}_{t,h}$. As with equations (1) and (2), we again cluster standard errors by survey date to account for possible correlations in the error terms associated with the two forecasting horizons at each time period.

2.4.1 Dispersion in Medians

We present the results for PGDP inflation in Table 2.3 and for RGDP growth in Table 2.4, with five versions of each equation, a baseline (a) without spread, macro uncertainty, and policy uncertainty, a second (b) including spread, and a third version including all variables (c). For the full set of variables, we run the regression for all horizons (c), and then splitting the sample into that for the shorter horizons (d) and the longer horizons (e). Splitting the sample allows us to analyze the behavior of dispersion of short-horizon and 'long-horizon' forecasts.

The key results from these regressions is that the coefficients on macroeconomic uncertainty are significant and positive, while forecast uncertainty plays little role in explaining dispersion of medians. The estimates in column (c) in both Tables 2.3 and 2.4 show that for both variables, dispersion is positively correlated with direct measures of macroeconomic (and policy) uncertainty, even after controlling for forecast horizon, lagged dispersion, and other variables. The evidence for the macroeconomic uncertainty index is more convincing than that for the policy uncertainty index, the coefficients on which are insignificant in columns (d) and (e) for both variables. The coefficient on *MU* is larger in the 'long-horizon' regression than in the short horizon regression, where the coefficient is not significant. This is particularly interesting as the macro-uncertainty index that we use measures predictability of a 3-month horizon. This suggests that dispersion is correlated to prevailing 'spot' levels of macroeconomic uncertainty, more so in the long-horizon where presumably there is less information of relevance to the forecasted variable. The coefficient in the short-horizon regression is not significant, though the value is still reasonably large. For PGDP inflation, the coefficients on average levels of forecast uncertainty are largely insignificant (they are only mildly significant in columns (d) and (e) of Table 2.3, the latter of the 'wrong' sign).

While the key results of interest are regarding the uncertainty indices, Table 2.3 also contain a number of other interesting results. The regression result for PGDP inflation in Table 2.3 show that lagged dispersion is large and significant, indicating that there is persistence in the level of dispersion from survey to survey, and this is true after controlling for horizons. This result is different from the persistence in the relative 'optimism' and 'pessimism' of individual forecasters (e.g., as observed in Consensus Economics forecast data by Patton and Timmermann 2010) which has more to do with relative rankings of the forecasters; our result says that the overall degree of dispersion is persistent. The coefficient becomes smaller as more explanatory variables are included, first spread then the uncertainty indices. It is interesting that the coefficient on lagged dispersion is larger for the short horizon forecasts than the long-horizon ones (comparing columns (d) and (e)) which may say something regarding the way the forecasters process information. The regression result for RGDP growth in Table 2.4 show similar results, with some important differences. Again, lagged dispersion in density forecast medians is large, positive, and significant, decaying as more explanatory variables are included. However, this coefficient is much smaller (and not significant) in the short-horizon regression, whereas it remains large and significant in the long-horizon regression. This may indicate more information processing in the shorter horizons as new information arrives.

The coefficient estimates on the horizon dummies also show some interesting patterns. In column (a) of Table 2.3, the horizon dummies imply decreasing dispersion with decreasing horizon, though less so in the regressions with the uncertainty indices. The decreasing dispersion with decreasing horizon is also less clear in for output growth, except perhaps for horizons 0 and 1. However, there does appear to be a spike in the dispersion at horizons 3 and 7, which corresponds to forecasts made in the first survey of the year. The spike in dispersion at the start of each year may suggest that views and information tend to be reevaluated or incorporated at the start of the year. These findings may support information-rigidity type explanations for dispersion, or it may be that annual-frequency variables (or annual-frequency versions of variables) are taken into account at the start of the year.

For PGDP inflation, the coefficient on the dispersion of forecasts on yield-spread is significant, i.e., the dispersion in forecasters' views regarding inflation is positively correlated to their dispersed views regarding spread. The coefficient on spread remains significant, though smaller, after inclusion of the uncertainty indices. Similarly, for RGDP growth, the coefficient of the T-bill rate forecast dispersion is positive and significant in the short-horizon forecasts. (The coefficient of the spread forecast dispersion is much weaker, when we replace the T-bill rate dispersion with the dispersion in spread), which is consistent with the Ang, Piazzesi, and Wei (2006) evidence that short-term interest rates forecast output growth better than spread.

As for the average levels of density forecast medians and skewness, they are by and large insignificant across Tables 2.3 and 2.4. The exception appears to be for overall level of the density forecast median. In regressions (a) and (b) of Table 2.4, dispersion in density forecast medians are negatively correlated with the overall expected levels of growth. As noted in Patton and Timmermann (2010), this is a result consistent with macroeconomic models that incorporate heterogeneous information. However, this correlation vanishes when macro and policy uncertainty are included in the regression.

2.4.2 Dispersion in Range and Skewness

The regressions for the dispersion of individual density forecast ranges are given in Table 2.5. The horizon dummies show that the forecasters disagree more on uncertainty as the forecast horizon declines. This is quite different from what was found for forecaster uncertainty and disagreement in medians. With the arrival of new information, forecasters adjust the shape of their forecasts and become more heterogeneous. This is generally true for both variables. For inflation, there appears to be very little persistence in the overall level of forecaster uncertainty, whereas the coefficient on lagged range dispersion is much stronger in the short-horizon regressions for output growth. The coefficients on the average level of forecast uncertainty is significant and positive, i.e., there is more disagreement in individual uncertainty when the average level of uncertainty is high. This indicates that there are some forecasters who tend always to report low uncertainty, even as others are reporting high uncertainty. More interesting is that dispersion in inflation and output growth forecaster uncertainty do not respond to policy uncertainty at all, or macro uncertainty in the short horizon, and respond differently to macroeconomic uncertainty for long-horizon forecasts. Dispersion in inflation forecaster uncertainty rises as macroeconomic uncertainty, indicating that the degree of heterogeneity in perceived uncertainty increases during more volatile periods. However, for output growth the effect is strongly and negatively significant, which may suggest that in periods where substantial uncertainty in real output growth is detected, forecasts update their perceived uncertainty to a higher level consistently and thus demonstrate a lower disagreement in forecaster uncertainty.

Table 2.6 shows the regressions for dispersion in skewness. The horizon dummies are all insignificant, and are left out of the output. Columns (a) and (d) are regressions on the full-sample, whereas columns (b) and (e) are the short-horizon forecasts, and (c) and (f) are the long-horizon forecasts. It does not appear that any of our variables can explain variations in the dispersion of individual density forecast skewness. There is no response to either policy and macroeconomic uncertainty for both variables. There are also no significant effects from either the dispersion of yield spreads or T-bill rates. There are some significant results regarding overall levels of skewness, but these are harder to interpret. While there are interesting results explaining the behavior of overall levels of skewness (Table 2.2), it seems there is no accounting for why the forecaster differ in the shape of their forecast densities.

2.4.3 Robustness Check

In this section, we conducted three robustness analyses to test the validity of our results. First, as shown in Figure 2.5, there is a spike in macroeconomic uncertainty during the crisis period in 2009. We would like to check whether the key findings are still present in the results and some confusing patterns potentially resulted from the spike in macroeconomic uncertainty would disappear when the crisis period is excluded. Second, we further examine the components of macroeconomic uncertainty, financial uncertainty and real uncertainty, to see which elements really matter in explaining dispersion. Finally, we add the components of policy uncertainty to check the robustness of our results and identify potential differences in their impact on disagreement.

A Excluding Crisis Period

Most of the results in the robustness check are consistent with the main analysis in the robustness check except for two important differences. First, after excluding crisis period, we find that macroeconomic uncertainty has quite limited effect in dispersion of density forecast medians of inflation. This shows that dispersion in forecasts for inflation are mainly affected by policy uncertainty and only responses to macroeconomic uncertainty in the very volatile period. Second, the surprising pattern found in Table 2.1 for average forecaster uncertainty of output growth, that PU and MU have significant and opposite effect, disappears in the regression excluding the crisis period. Table 2.8 compares the results of the full sample (shown in left three columns) and excluding-crisis sample (shown in right three columns). The regression with full sample shows that forecaster uncertainty is positively correlated with PU but negatively related to MU, which is quite confusing and we attribute it to the spike of MU during the crisis in 2009. The right panel verifies this explanation as we no longer observe any significant relationship between forecaster uncertainty and the other two uncertainty measures. This confirms our conclusion that forecaster uncertainty is hardly correlated with either PU or MU.

B Components of Macroeconomic Uncertainty

JLN exploit two datasets to form forecasts and construct objective measures of uncertainty. The first dataset, denoted X_m, consists of 132 (134 in the latest version) mostly macroeconomic series. The other one, denoted X_f, consists of 147 (148 in the latest version) financial time series. They combine the macro and financial datasets together to estimate forecasting factors in these 279 series, but estimate macroeconomic uncertainty for the individual uncertainties in the 132 macro series only. This means that the macro uncertainty measure we use excludes the financial series. To further analyze which component of the macro uncertainty contributes more to our findings, we refer to another related study of Ludvigson, Ma and Ng (2015), which explicitly distinguishes macro uncertainty from financial uncertainty by constructing a measure of uncertainty based on the 147 financial time series in X_f and obtains the real uncertainty based on real activity variables in X_m. Table 2.7 summarizes the major difference of the three uncertainty measures.

The robustness tests show that macro uncertainty (MU) and real uncertainty (RU) have similar power in explaining dispersion among medians of density forecasts for both inflation and output growth. However, when we substitute financial uncertainty ($FinU$) into the regression, the results are quite different for the two variables. $FinU$ has little effect on PGDP inflation, but shows even more strong correlation with RGDP growth.

The robustness tests for dispersion in range and skewness give similar conclusion to those in Table 2.5 and Table 2.6 in terms of both significance and magnitude of the coefficients.

The robustness tests reinforce our findings that macro uncertainty plays an important role in determining dispersion of density forecasts. By further examining the role of components of macro uncertainty, we find that dispersion in medians of inflation forecasts is mainly affected by uncertainty in macro indicators while that of output growth forecasts is more sensitive to uncertainty in financial series.

C Components of Policy Uncertainty

Recently, the policy uncertainty index developed by Baker, Bloom and Davis is enriched by a set of elements: economic policy uncertainty, monetary policy, fiscal policy (taxes or spending), taxes, government spending, health care, national security, entitlement programs, regulation (and financial regulation), trade policy, sovereign debt and currency crises. Accordingly, we extend our robustness analyses to include all the categories to investigate which of them plays a more important role in explaining dispersion. The results suggest that, among all the components, national security and trade policy have little effect in explaining dispersion, while financial regulation has strong effect that dominates MU, with the effect of MU disappearing when including financial regulation in the regression. The rest of the factors show patterns that are similar to our main analysis. Overall the result in this part is consistent with our main finding that dispersion is strongly correlated to the macro and policy uncertainty.

2.5 Concluding remarks

We explore the role of uncertainty in explaining dispersion in the median, central 90% interval, and a skewness measure of professional forecasters' density forecasts of real output growth and inflation. We consider three separate notions of uncertainty: an objective measure of macroeconomic uncertainty capturing the fact that macroeconomic variables are easier to forecast at some times than at others, policy uncertainty, and forecaster uncertainty.

The empirical evidence suggests that dispersion among forecasters in medians is related to overall macroeconomic uncertainty, and to a lesser extent policy uncertainty. The dispersion shows a different pattern in short horizon versus longer horizon. In longer horizon, we observe a larger but smoother degree of disagreement which relies more on past information (e.g., lagged dispersion) and overall macroeconomic uncertainty index while the degree of dispersion is more related to controls closely linked to new information such as dispersion in interest rates in short horizons.

The link between macroeconomic uncertainty and forecast dispersion appears to apply mainly to the location of the density forecasts, and is much weaker for dispersion in forecaster uncertainty. As for the dispersion in forecaster uncertainty, our results provide evidence that professional forecasters are heterogeneous in the way they revise their subjective forecast distributions: the dispersion in forecaster uncertainty rises as forecast horizon shortens, and exhibits a positive correlation with average forecaster uncertainty, suggesting that some forecasters tend not to revise their reported levels of uncertainty upwards. Dispersion in forecaster uncertainty is positively correlated with macroeconomic economy uncertainty only in the long run, though in opposite directions for inflation and for real output forecasters. In the latter case, forecasters disagree less on perceived uncertainty when there is an increase in real output uncertainty. For skewness there appears no link whatsoever between dispersion and macro, policy, or forecaster uncertainty (or any of our other controls).

In order to test for the validity of our results, we conducted several robustness analyses from three dimensions. First, we check the results using sample excluding this period to see whether the key findings are still present in the results and some confusing patterns potentially resulted from the spike in macroeconomic uncertainty would disappear. The results confirm our conclusion that forecaster uncertainty is hardly correlated with either PU or MU. Second, we examine the components of macroeconomic uncertainty, financial uncertainty and real uncertainty, to see which elements really matter in explaining dispersion. The tests reinforce our findings that macro uncertainty plays an important role in determining dispersion of density forecasts. By further examining the components of macro uncertainty, we find that dispersion in medians of inflation forecasts is mainly affected by uncertainty in macro indicators while that of output growth forecasts is more sensitive to uncertainty in financial series. Finally, we add the components of policy uncertainty to check the robustness of our results and identify potential differences in their impact on disagreement. The results suggest that, among all the components, national security and trade policy have little effect in explaining dispersion, while financial regulation has strong effect that dominates MU. The rest of the factors show similar results to our main analysis. The analysis in this part is consistent with our main results that dispersion is strongly correlated to the macro and policy uncertainty.

Overall, average forecaster uncertainty appears to have little role in explaining forecast dispersion. Furthermore, there is evidence that forecaster uncertainty appears to be affected by survey design and there appears a need for further research into forecaster uncertainty elicitation methods.

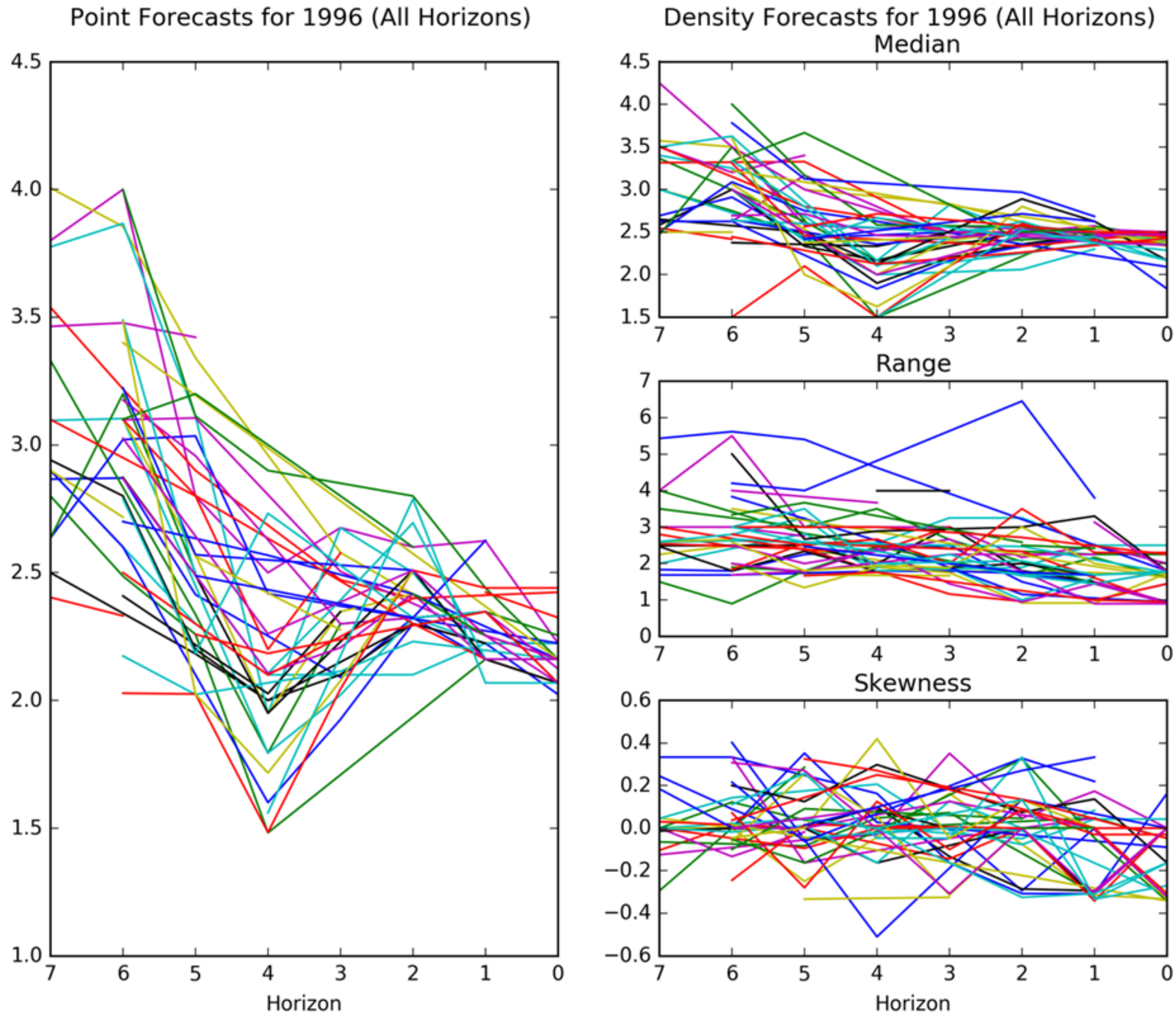

Figure 2.1 Point Forecasts and Density Forecasts of SPF

Left diagram: point forecasts of all forecasters for 1996 PGDP inflation from 8 surveys. Horizons 7 to 4 forecasts were made in the 1995Q1-Q4 surveys, and horizons 3 to 0 forecasts were made in the 1996Q1-Q4 surveys. Right column shows the three key characteristics (median, range and skewness) of density forecasts of all forecasters for 1996 PGDP inflation from the same 8 surveys.

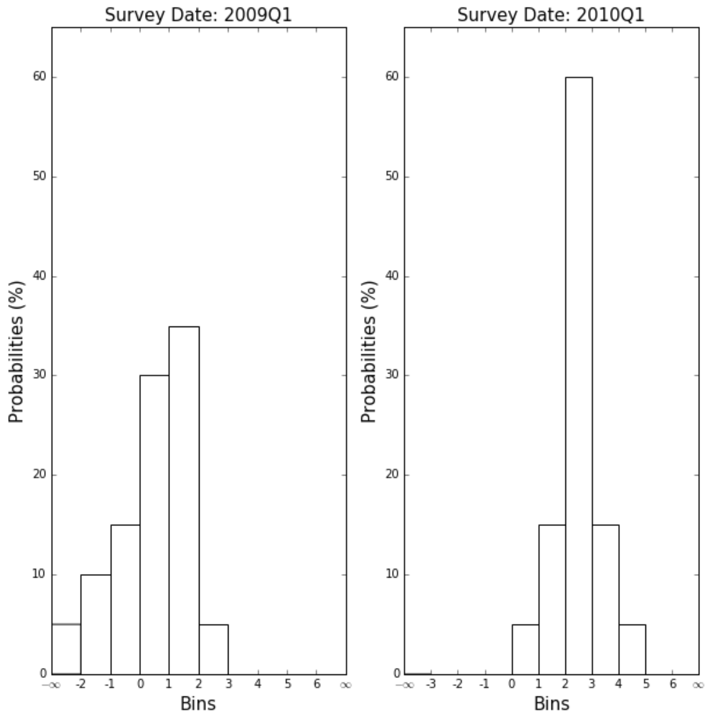

Figure 2.2 Density Forecasts: Example

Density forecasts of annual average RGDP growth in 2010 made by Forecaster 516 in the 2009Q1 and 2010Q1 surveys.

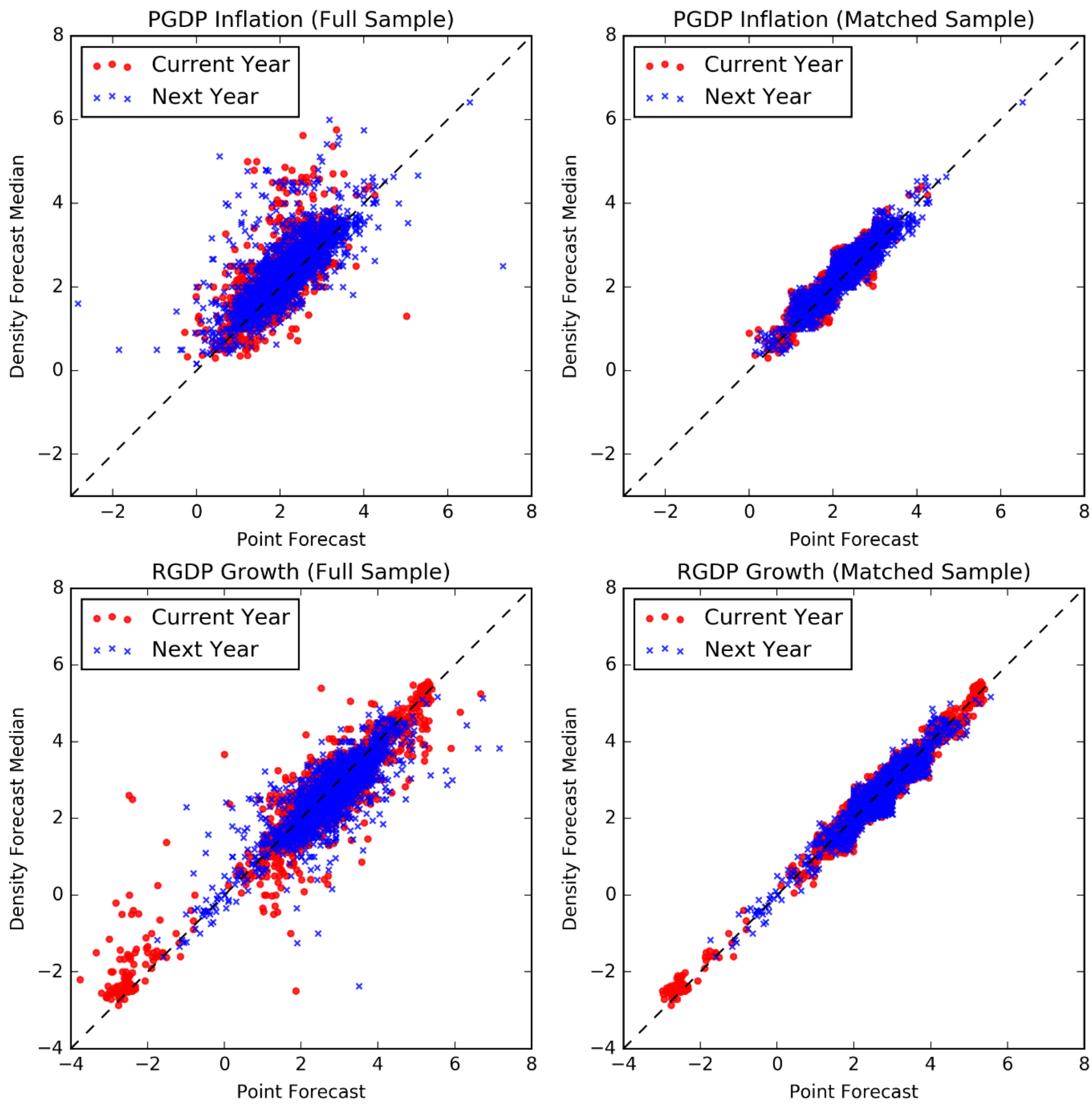

Figure 2.3 Density Forecasts vs Point Forecasts

Plots of density forecast medians against forecasters corresponding point forecast, full and matched samples.

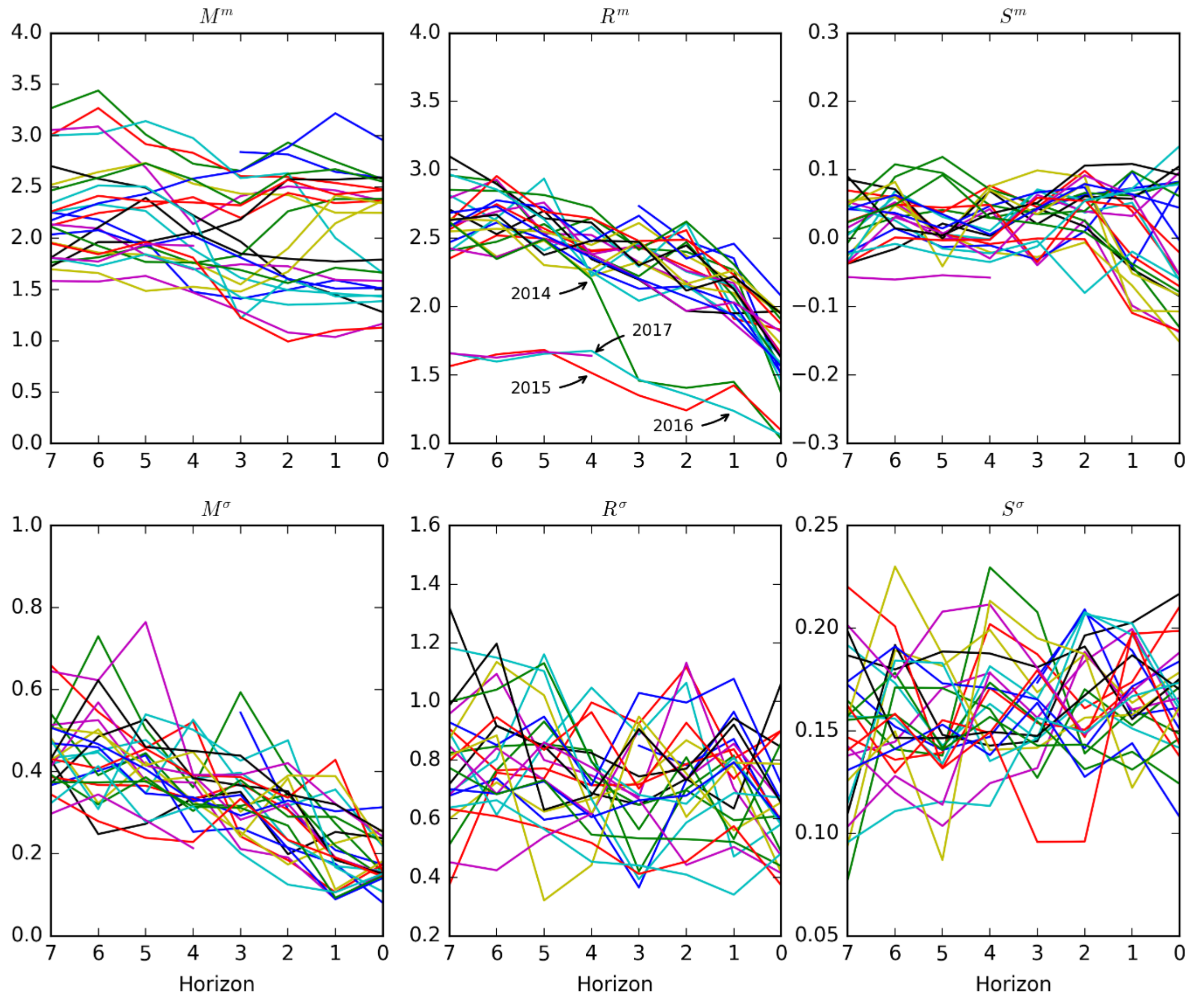

Figure 2.4 Features of Density Forecasts

Figure 2.4(a) PGDP. Top row: average of individual PGDP inflation density forecast medians, range, and skewness. Bottom row: standard deviation of individual PGDP inflation density forecast medians, range, and skewness, depicting density forecast dispersion. Each line corresponds to a target year, all plotted against forecast horizon.

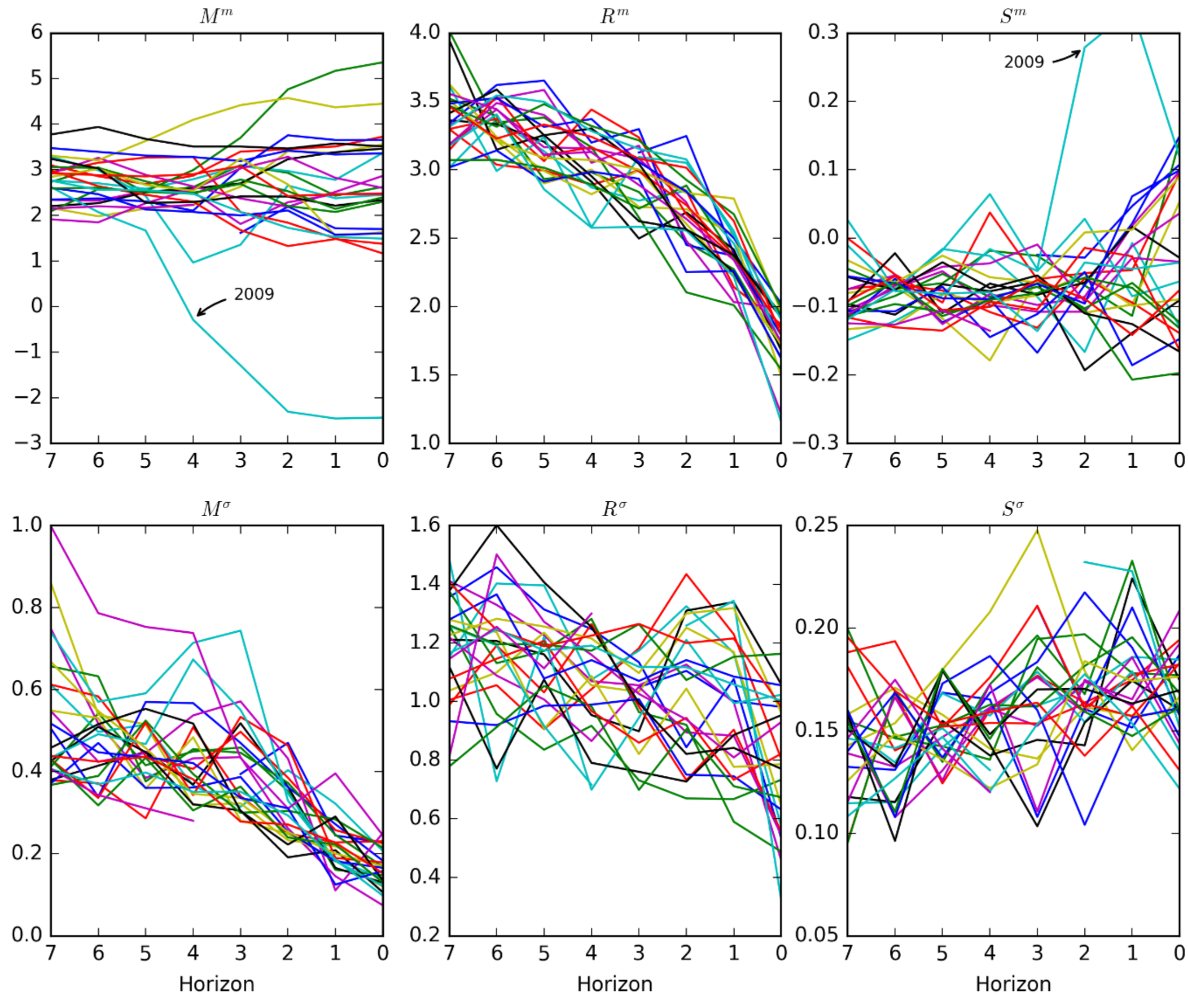

Figure 2.4(b) RGDP. Top row: average of individual RGDP growth density forecast medians, range, and skewness. Bottom row: standard deviation of individual RGDP growth density forecast medians, range, and skewness, depicting density forecast dispersion. Each line corresponds to a target year, all plotted against forecast horizon.

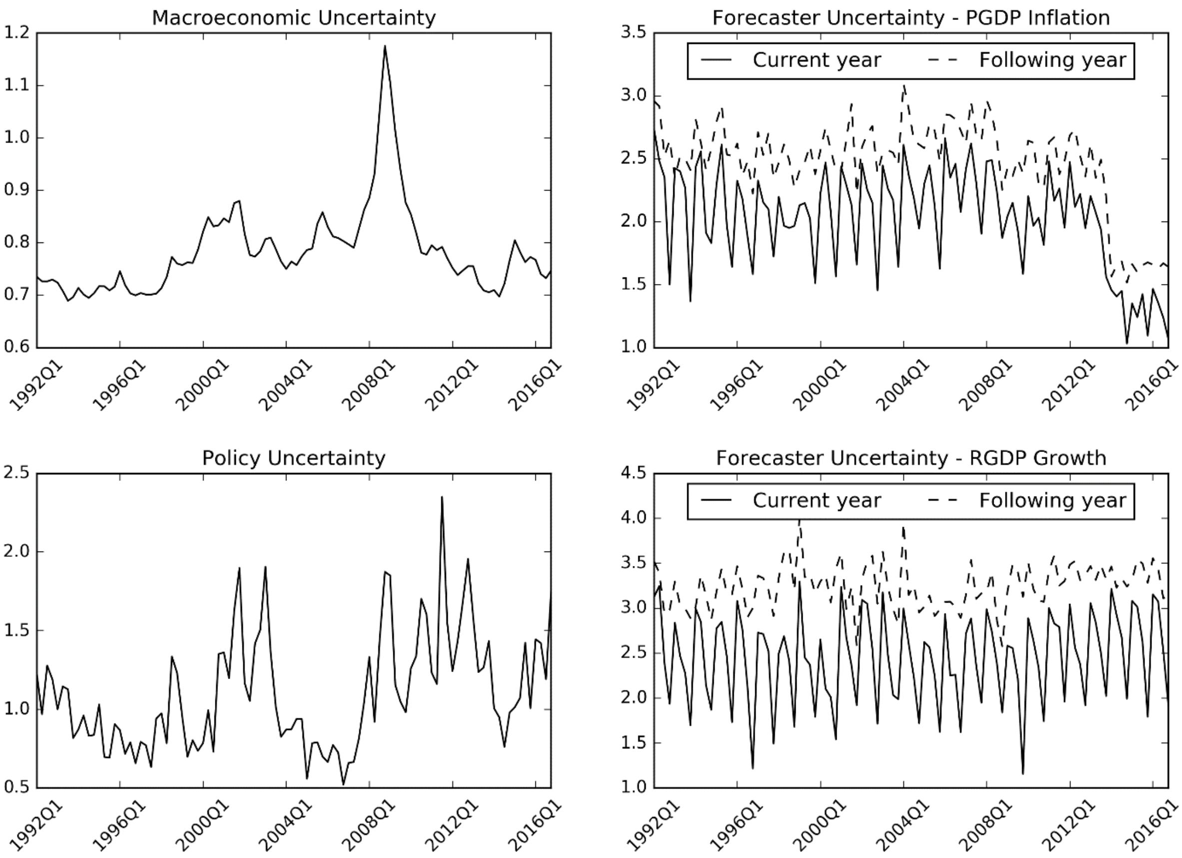

Figure 2.5 Uncertainties

Top left displays the time series plot of macroeconomic uncertainty MU_t. Bottom left shows the time series plot of policy uncertainty PU_t. Diagrams in the right column show forecast uncertainty for PGDP Inflation and RGDP growth corresponding to the current and following year forecasts made at each survey.

Table 2.1 Average Forecaster Uncertainty

Variable	PGDP Inflation			RGDP Growth		
	(a)	(b)	(c)	(d)	(e)	(f)
h_1	0.376***	0.394***		0.536***	0.520***	
	(7.18)	(7.85)		(10.08)	(10.17)	
h_2	0.513***	0.552***		0.841***	0.820***	
	(9.47)	(9.92)		(11.81)	(11.14)	
h_3	0.653***	0.617***		1.340***	1.373***	
	(11.86)	(11.26)		(17.69)	(15.69)	
h_4	0.629***			1.121***		
	(11.48)			(14.82)		
h_5	0.753***		0.123***	1.222***		0.107**
	(12.17)		(3.16)	(15.96)		(2.10)
h_6	0.819***		0.181***	1.369***		0.249***
	(12.47)		(4.42)	(16.49)		(4.77)
h_7	0.821***		0.200***	1.489***		0.351***
	(13.87)		(3.87)	(19.43)		(5.62)
Newbin	-0.723***	-0.716***	-0.735***	0.089**	0.061	0.119**
	(-10.00)	(-10.12)	(-7.30)	(2.31)	(1.31)	(2.31)
lagged R^m	0.103	-0.020	0.182*	0.224***	0.273***	0.156
	(1.41)	(-0.25)	(1.83)	(3.32)	(2.94)	(1.60)
M^m	0.118***	0.179***	0.069	0.004	0.001	0.024
	(3.90)	(5.03)	(1.49)	(0.19)	(0.06)	(0.48)
PU	-0.024	0.066	-0.106**	0.101**	0.153**	0.055
	(-0.66)	(1.52)	(-2.44)	(2.03)	(2.32)	(0.81)
MU	0.194	0.113	0.225	-0.467**	-0.512*	-0.411*
	(1.12)	(0.61)	(1.01)	(-2.53)	(-1.68)	(-1.69)
Constant	1.155***	1.241***	1.764***	1.434***	1.310***	2.722***
	(5.86)	(6.11)	(6.07)	(4.91)	(2.98)	(6.82)
Obs.	198	99	99	198	99	99
R^2	0.88	0.86	0.84	0.88	0.85	0.40
r(1)	1.326	1.317		1.374	1.397	
r(2)	1.089	1.097		1.155	1.164	
r(3)	1.084	1.036		1.219	1.260	
r(4)	0.987			0.921		
r(5)	1.070		1.070	1.040		1.039
r(6)	1.035		1.031	1.055		1.050
r(7)	1.001		1.010	1.043		1.034

Notes: Regression results for average individual uncertainty R^m, t-statistics in parentheses, calculated from standard errors clustered by year and quarter. Regressions (a) and (d) for full sample, regressions (b) and (e) for 'current year' forecasts, and regressions (c) and (f) for next year forecasts. The entries r(k) are the ratios of the average range at horizon k to horizon $k-1$.

Table 2.2 Average Individual Skewness

Variable	PGDP Inflation			RGDP Growth		
	(a)	(b)	(c)	(d)	(e)	(f)
Newbin	0.005	0.013	-0.009	0.013	0.010	0.013
	(0.22)	(0.37)	(-0.48)	(1.16)	(0.53)	(1.62)
lagged S^m	0.392***	0.403***	0.271**	0.323***	0.292***	0.319***
	(4.46)	(3.70)	(2.49)	(4.04)	(3.14)	(2.83)
M^m	-0.022**	-0.022	-0.015	-0.029***	-0.036***	-0.011
	(-2.26)	(-1.50)	(-1.61)	(-3.66)	(-3.84)	(-1.35)
R^m	0.043*	0.035	0.048**	-0.034	-0.015	-0.054***
	(1.79)	(0.82)	(2.44)	(-1.53)	(-0.39)	(-3.20)
PU	-0.012	-0.012	-0.011	-0.029*	-0.051*	-0.005
	(-1.06)	(-0.63)	(-1.23)	(-1.72)	(-1.89)	(-0.42)
MU	-0.086**	-0.106	-0.059*	-0.039	-0.103	0.018
	(-2.28)	(-1.52)	(-1.88)	(-0.58)	(-0.87)	(0.29)
Obs	198	99	99	198	99	99
R^2	0.34	0.31	0.46	0.40	0.44	0.25

Notes: Results for average skewness S^m regressions. Robust t-statistics in parentheses, calculated from standard errors clustered by year and quarter. Constant and horizon dummies included but omitted from the table. Regressions (a) and (d) for full sample, regressions (b) and (e) for 'current year' forecasts, and regressions (c) and (f) for next year forecasts.

Table 2.3 Dispersion of Individual PGDP Inflation Density Forecast Medians

	Dep. Var. PGDP Inflation Density Forecast Medians M^m				
	(a)	(b)	(c)	(d)	(e)
h_1	0.005	-0.001	0.009	-0.026	
	(0.20)	(-0.04)	(0.34)	(-0.94)	
h_2	0.060*	0.046	0.065**	0.016	
	(1.84)	(1.32)	(1.99)	(0.47)	
h_3	0.157***	0.123***	0.132***	0.076**	
	(4.39)	(3.44)	(3.94)	(2.04)	
h_4	0.107***	0.066	0.090**		
	(2.84)	(1.59)	(2.33)		
h_5	0.145***	0.097**	0.126***		0.048**
	(3.87)	(2.28)	(3.09)		(2.25)
h_6	0.170***	0.110**	0.145***		0.072***
	(3.71)	(2.10)	(2.99)		(2.81)
h_7	0.197***	0.118**	0.149***		0.065***
	(4.51)	(2.42)	(3.18)		(2.71)
newbin	-0.055	-0.058	-0.057	0.023	-0.179***
	(-1.48)	(-1.53)	(-1.63)	(0.68)	(-2.90)
lagged M^δ	0.334***	0.288***	0.225***	0.231**	0.157*
	(5.47)	(4.85)	(3.85)	(2.57)	(1.68)
M^m	0.000	0.009	0.027	0.006	0.052*
	(0.01)	(0.52)	(1.49)	(0.33)	(1.91)
R^m	0.015	0.003	-0.006	0.081*	-0.112*
	(0.36)	(0.07)	(-0.15)	(1.94)	(-1.69)
S^m	0.082	0.065	0.167	0.149	0.064
	(0.73)	(0.60)	(1.65)	(1.31)	(0.23)
spread		0.219***	0.178***	0.193**	0.220***
		(3.37)	(2.83)	(2.07)	(3.21)
PU			0.036**	0.032	0.025
			(2.22)	(1.42)	(0.89)
MU			0.184*	0.135	0.259**
			(1.85)	(0.93)	(2.47)
constant	0.090	0.086	-0.104	-0.174	0.165
	(1.37)	(1.29)	(-1.01)	(-1.37)	(0.84)
Obs	198	198	198	99	99
R^2	0.64	0.66	0.69	0.56	0.48

Notes: Results for dispersion of individual PGDP inflation density forecast medians (M^σ) regressions. Robust t-statistics in parentheses, calculated from standard errors clustered by year and quarter.

Table 2.4 Dispersion of Individual RGDP Growth Density Forecast Medians

	Dep. Var. RGDP Growth Density Forecast Medians M^m				
	(a)	(b)	(c)	(d)	(e)
h_1	-0.020	-0.019	-0.013	-0.005	
	(-0.61)	(-0.57)	(-0.39)	(-0.15)	
h_2	0.026	0.023	0.039	0.046	
	(0.68)	(0.59)	(0.93)	(0.93)	
h_3	0.215^{***}	0.194^{***}	0.192^{***}	0.138^{**}	
	(4.65)	(3.83)	(3.84)	(2.47)	
h_4	0.113^{*}	0.089	0.105^{*}		
	(1.95)	(1.49)	(1.76)		
h_5	0.117^{**}	0.081	0.103		0.007
	(2.03)	(1.29)	(1.62)		(0.26)
h_6	0.107^{*}	0.062	0.094		0.003
	(1.72)	(0.90)	(1.34)		(0.10)
h_7	0.205^{***}	0.151^{*}	0.173^{**}		0.086^{**}
	(2.90)	(1.95)	(2.27)		(2.33)
newbin	-0.044^{***}	-0.033^{*}	-0.046^{***}	-0.041^{*}	-0.047^{*}
	(-2.79)	(-1.96)	(-2.67)	(-1.97)	(-1.88)
lagged M^{δ}	0.441^{***}	0.394^{***}	0.343^{***}	0.141	0.377^{***}
	(5.08)	(4.26)	(4.42)	(1.36)	(4.09)
M^m	-0.021^{**}	-0.022^{**}	-0.008	-0.010	-0.033
	(-2.35)	(-2.54)	(-0.94)	(-1.32)	(-1.38)
R^m	0.041	0.039	0.051	0.055	0.050
	(1.09)	(1.02)	(1.37)	(1.40)	(0.91)
S^m	-0.046	-0.064	0.035	0.008	0.196
	(-0.44)	(-0.60)	(0.34)	(0.10)	(0.61)
TBill		0.134^{**}	0.086	0.322^{**}	0.045
		(2.13)	(1.29)	(2.01)	(0.52)
PU			0.045^{**}	0.038	0.040
			(2.41)	(1.57)	(1.29)
MU			0.233^{**}	0.019	0.387^{**}
			(2.06)	(0.17)	(2.30)
constant	0.065	0.070	-0.204	0.001	-0.136
	(0.83)	(0.89)	(-1.53)	(0.01)	(-0.59)
Obs	198	198	198	99	99
R^2	0.70	0.71	0.73	0.71	0.51

Notes: Results for dispersion of individual RGDP growth density forecast medians (M^{σ}) regressions. Robust t-statistics in parentheses, calculated from standard errors clustered by year and quarter.

Table 2.5 Dispersion of Individual Density Forecast Range

Variable	PGDP Inflation			RGDP Growth		
	(a)	(b)	(c)	(d)	(e)	(f)
h_1	-0.135**	-0.093		-0.045	-0.061	
	(-2.38)	(-1.43)		(-0.71)	(-0.79)	
h_2	-0.207***	-0.141*		-0.138*	-0.153	
	(-3.17)	(-1.83)		(-1.72)	(-1.48)	
h_3	-0.297***	-0.212**		-0.206**	-0.199	
	(-3.94)	(-2.36)		(-2.12)	(-1.52)	
h_4	-0.340***			-0.272***		
	(-4.86)			(-2.87)		
h_5	-0.344***		-0.013	-0.280***		-0.005
	(-4.07)		(-0.37)	(-2.69)		(-0.13)
h_6	-0.360***		-0.044	-0.302***		-0.029
	(-3.98)		(-0.93)	(-2.68)		(-0.59)
h_7	-0.381***		-0.065	-0.286**		-0.029
	(-4.02)		(-1.36)	(-2.40)		(-0.51)
Newbin	0.239***	0.114	0.374***	0.076**	0.045	0.118***
	(3.57)	(1.50)	(3.85)	(2.38)	(0.84)	(2.90)
lagged R^{δ}	0.081	0.033	0.039	0.225***	0.315***	0.078
	(1.27)	(0.35)	(0.44)	(2.96)	(3.31)	(0.70)
M^m	0.023	0.033	0.028	0.003	0.016	-0.047
	(0.67)	(0.72)	(0.62)	(0.14)	(0.65)	(-1.37)
R^m	0.501***	0.435***	0.578***	0.389***	0.418***	0.374***
	(6.19)	(4.06)	(4.75)	(5.49)	(4.32)	(3.83)
S^m	0.300	0.296	0.479	0.073	0.171	-0.188
	(1.31)	(1.12)	(0.91)	(0.31)	(0.56)	(-0.45)
Spread/TBill	0.170	-0.075	0.195	0.086	-0.025	0.170*
	(1.41)	(-0.38)	(1.53)	(0.91)	(-0.07)	(1.71)
PU	-0.039	-0.015	-0.050	-0.017	-0.005	-0.040
	(-1.38)	(-0.41)	(-1.48)	(-0.44)	(-0.09)	(-0.78)
MU	0.196	-0.019	0.472***	-0.238	-0.149	-0.376*
	(1.63)	(-0.09)	(2.71)	(-1.30)	(-0.48)	(-1.83)
Constant	-0.431***	-0.134	-1.159***	-0.009	0.047	-0.193
	(-2.86)	(-0.66)	(-4.31)	(0.20)	(-0.59)	(0.51)
Obs	198	99	99	196	97	99
R^2	0.55	0.50	0.62	0.59	0.55	0.49

Notes: Results for dispersion of individual RGDP growth density forecast range (R^{σ}) regressions. Robust t-statistics in parentheses, calculated from standard errors clustered by year and quarter. Dispersion of yield spread forecasts (*spread*) used for PGDP regressions, dispersion of T-Bill rate forecasts (TBill) used for RGDP regressions.

Table 2.6 Dispersion of Individual Density Forecast Skewness

Variable	PGDP			RGDP		
	(a)	(b)	(c)	(d)	(e)	(f)
Newbin	-0.017	-0.005	-0.038	-0.007	-0.017*	0.002
	(-1.24)	(-0.35)	(-1.65)	(-1.36)	(-1.81)	(0.33)
lagged S^{δ}	0.103	0.113	0.039	0.084	0.136	-0.039
	(1.28)	(1.05)	(0.33)	(1.27)	(1.51)	(-0.38)
M^m	-0.006	-0.012*	0.002	0.004	0.004	0.009*
	(-1.35)	(-1.90)	(0.28)	(1.56)	(1.17)	(1.80)
R^m	0.011	0.027*	-0.023	0.004	0.010	-0.014
	(0.86)	(1.77)	(-1.09)	(0.37)	(0.72)	(-1.05)
S^m	-0.033	-0.084**	0.132	0.031	0.072**	-0.137**
	(-1.00)	(-2.44)	(1.27)	(0.94)	(2.05)	(-2.01)
Spread/TBill	0.005	-0.071*	0.029	-0.026	-0.059	-0.021
	(0.19)	(-1.82)	(0.97)	(-1.35)	(-0.78)	(-1.19)
PU	-0.011*	-0.008	-0.017	0.000	0.006	-0.001
	(-1.79)	(-0.95)	(-1.64)	(0.05)	(0.85)	(-0.12)
MU	0.030	0.019	0.061	0.040	0.055	0.044
	(1.28)	(0.55)	(1.24)	(1.62)	(1.44)	(1.27)
Constant	0.132***	0.123***	0.173**	0.106***	0.075	0.142***
	(3.94)	(2.90)	(2.59)	(3.17)	(1.58)	(2.70)
Obs	198	99	99	196	97	99
R^2	0.20	0.22	0.25	0.22	0.16	0.17

Notes: Results for dispersion of individual RGDP growth density forecast skewness (S^{σ}) regressions. Robust t-statistics in parentheses, calculated from standard errors clustered by year and quarter. Dispersion of yield spread forecasts (*spread*) used for PGDP regressions, dispersion of T-Bill rate forecasts (TBill) used for RGDP regressions. The horizon dummies are included in the regression but are not included in the table. Columns (a) and (d) are regressions on the full-sample, whereas columns (b) and (e) are the short-horizon forecasts, and (c) and (f) are the long-horizon forecasts.

Table 2.7 Components of *MU*

Category	Dataset	Number of series
(M): Macro	All series in X_m	134
(F): Financial	All series in X_f	148
(R): Real	Real activity series in X_m	73

Notes: A summary of three uncertainty measures related to MU. The macro uncertainty is constructed based on the 134 macroeconomic series. The financial uncertainty is measured based on the 147 financial time series in X_f. The real uncertainty is derived from real activity variables in X_m.

Table 2.8 Average Forecaster Uncertainty of RGDP Growth

Variable	Full			Excluding Crisis		
	(a)	(b)	(c)	(d)	(e)	(f)
h_1	0.536***	0.520***		0.532***	0.534***	
	(10.08)	(10.17)		(11.39)	(10.96)	
h_2	0.841***	0.820***		0.854***	0.860***	
	(11.81)	(11.14)		(12.59)	(12.49)	
h_3	1.340***	1.373***		1.299***	1.289***	
	(17.69)	(15.69)		(17.87)	(14.19)	
h_4	1.121***			1.148***		
	(14.82)			(16.46)		
h_5	1.222***		0.107**	1.250***		0.100*
	(15.96)		(2.10)	(18.03)		(1.91)
h_6	1.369***		0.249***	1.384***		0.229***
	(16.49)		(4.77)	(17.27)		(4.22)
h_7	1.489***		0.351***	1.515***		0.357***
	(19.43)		(5.62)	(19.84)		(5.56)
Newbin	0.089**	0.061	0.119**	0.085*	0.087**	0.081
	(2.31)	(1.31)	(2.31)	(1.94)	(2.04)	(1.40)
lagged R^m	0.224***	0.273***	0.156	0.174**	0.161	0.166*
	(3.32)	(2.94)	(1.60)	(2.63)	(1.62)	(1.72)
M^m	0.004	0.001	0.024	-0.059**	-0.082***	0.004
	(0.19)	(0.06)	(0.48)	(-2.24)	(-3.02)	(0.05)
PU	0.101**	0.153**	0.055	0.066	0.046	0.090
	(2.03)	(2.32)	(0.81)	(1.13)	(0.69)	(1.23)
MU	-0.467**	-0.512*	-0.411*	-0.183	-0.194	-0.301
	(-2.53)	(-1.68)	(-1.69)	(-0.52)	(-0.45)	(-0.58)
Constant	1.434***	1.310***	2.722***	1.569***	1.694***	2.639***
	(4.91)	(2.98)	(6.82)	(4.48)	(3.16)	(5.96)
Obs.	198	99	99	182	91	91
R^2	0.88	0.85	0.40	0.89	0.87	0.37

Notes: Regression results for average individual uncertainty R^m of output growth, t-statistics in parentheses, calculated from standard errors clustered by year and quarter. Regressions (a) – (c) for full sample, regressions (d) - (f) for the sample excluding crisis (1992-2008 and 2010-2017).

Chapter 3 Survey Design and Forecasters' Behavior: Evidence from Survey of Professional Forecasters

3.1 Introduction

Density forecasts are nowadays commonly included in professional surveys to depict a more complete picture of forecast uncertainty. In particular, they provide a direct self-represented measure of forecaster uncertainty, which is usually defined in the literature as the average of individual forecaster uncertainty, measured by inter-percentile range or standard deviation of each forecaster's density forecast. Surveys of density forecasts therefore serve as a rich resource for research on topics pertaining to forecaster uncertainty. Some of the studies focus on the features of forecaster uncertainty. Boero, Smith and Wallis (2015) find that there is great heterogeneity and persistence in forecaster uncertainty.

Uncertainty has always been the focus of policy makers but is difficulty to quantitatively measure. Thus dispersion, which measures the degree of disagreement among forecasters, has been widely taken as a substitute for uncertainty. As density forecasts make it possible to measure forecaster uncertainty, another set of studies utilize this direct measure to justify how well that dispersion can fit the movements of uncertainty (Rich and Tracy, 2008; Boero, Smith and Wallis (2008,2015). However, this relationship lacks theoretical support and remains a long-lasting debate in the literature.

Motivated by findings in Chapter two that forecaster uncertainty might be significantly affected by the change in survey design, we notice the importance of examining how the survey design influences the forecasters' behavior. In this chapter, we explore this question further by taking advantage of changes to the density forecast bins in three important surveys and examining the effect of such changes on forecaster uncertainty and dispersion among forecasters. By doing so, we are able to contribute to the two major strands of literature trying to investigate the features of forecaster uncertainty and the link between dispersion and uncertainty through a new dimension. If survey design has unanticipated and heterogeneous effects on dispersion and uncertainty, we can not only uncover the new features of forecaster uncertainty, but also better understand the confusing relationship between dispersion and uncertainty.

Moreover, our study provides implications on how survey design could affect forecasters' response. Modification of questionnaire in the survey is always costly and the designer would not change the survey unless when they have to. Unfortunately, there are always situations that force the surveyors to change the design. For example, when there is a sudden drop in the economy where the underlying indicator falls far beyond the lower bound of the bins, they must adjust the bins otherwise no valid information could be drawn from the forecasts: all of them are clustered in the lower open interval. There are also cases where the change is barely related to real economy: the original bins are poorly set and improvement is necessary to make the data more informative. Though such changes have been observed in the survey data, very few studies pay attention to whether these survey adjustments incur unexpected response of forecasters.

The surveys included in our sample are Survey of Professional Forecasters by the Philadelphia Fed (US-SPF), Survey of Professional Forecasters by the European Central Bank (ECB-SPF), and Survey of External Forecasters by Bank of England (SEF). They are the most well-known and widely studied surveys of professional forecasters. More importantly, they share many common features and thus make a clean setting for further analysis. For example, during the period of 2008-2009, the countries and regions covered in the data all suffered from a crisis - a sudden sharp drop in the real output growth. By looking into the different responses of both survey designers and forecasters, we are able to see how such interactions take place and affect forecasters' behaviors.

In some cases, adjustment of forecast bins reasonably arises from the fluctuation of underlying macroeconomic variable. Among the three surveys, ECB-SPF is more sensitive to the changes in the underlying indicators while SEF is most reluctant to do modifications. For example, during 2008-2009, as mentioned above, all the countries and regions in the sample experienced a big drop in output growth. The real value was so low that it fell beyond the lower bound of the interior intervals in all three surveys. After this surprising shock, ECB-SPF adjusted the ranges twice to completely include the full range of indictor's values in the inner intervals. However, SPF and SEF only slightly adjusted the intervals by 0.01 to 0.02. In other cases, however, the changes are neutral to the economic environment. It occurs when the current ranges are consistently larger than the variation of the real indicators, leading the survey designers to shrink the total range, sometimes together with the width of the intervals, to better fit the possible amplitude of the indicator's movement. This category of changes provides a cleaner setting that allows us to distinguish the framing effect from the effect of real economy. The first case where changes to survey design are necessitated by the variable (level or dispersion) is responding to changes in regimes (values of the underlying variables) while the second case where the changes are independent of the values of the variables is not prompted by any structural change of the underlying indicators. Hence, whether the second one has any effect on the forecasters behavior would be very interesting and provide important implications on survey designers who organize the data and analysist as well as policy makers who would draw reference from the surveys.

Indeed, the two different sets of adjustment in survey design discussed above have different impacts on the forecasters' behavior. We focus on dispersion and uncertainty specifically in this study as they are the most interesting measures to both researchers and policymakers. The dispersion, defined as the standard deviation of medians among different forecasters, only responds to the survey design changes triggered by real economy. While the forecaster uncertainty, measured by average of the 90 percent range of the individual density forecasts, responds to both cases. More importantly, the effects on uncertainty is more persistent than on dispersion.

Our work adds to the literature explaining why forecasters disagree from an additional dimension of survey design. Various explanations in the literature considering different types of heterogeneity among forecasters have been proposed to answer this question. The first one is heterogeneity in information set. This relates to information rigidity literature including noisy information model (Sims, 2002) where the dispersion stems from uncertain news, and sticky information model proposed by Mankiw and Reis (2002), where forecasters only occasionally update their information at different rates due to the cost in updating news. Coibion and Gorodnichenko (2012) provide empirical evidence to support the sticky information model. Another form of heterogeneity is linked to the models used by the forecasters. Branch (2004,2007) finds that forecasters dynamically switch between a set of costly alternative prediction models in each period. In addition to heterogeneity in models, Patton and Timmermann (2010) and Lahiri and Sheng (2010) show that difference in priors is an important determinant of dispersion, especially at longer horizons. Manzan (2011) extends the literature by illustrating that heterogeneity in interpretation of news as well could explain the source of dispersion. Last but not least, Capistran and Timmermann (2009) examine the role of loss function and find that heterogeneity in the asymmetric loss function could lead to dispersion. Our results provide evidence of the framing effect as potential source of bounded rationality and dispersion.

Our results also clarify the different features between dispersion and forecaster uncertainty, which contributes to another strand of literature that examine whether dispersion is a good measure for uncertainty. Since uncertainty is always the focus of policy makers as it plays an important role in the effect of any monetary policies and in absence of a well-accepted measure of uncertainty, dispersion is widely taken as a substitute for uncertainty. However, there is neither theoretical support nor consistent empirical evidence on the validity of this relationship. Zarnowitz and Lambros (1987) and Giordani and Soderlind (2003) find that the dispersion and uncertainty are highly correlated so that dispersion can serve as a reasonable proxy for uncertainty. On the contrary, Rich and Tracy (2008) argue that there is little evidence supporting this relationship. Furthermore, Boero, Smith and Wallis (2008, 2015) examine this topic using the Bank of England Survey of External Forecasters and argue that dispersion is not a good proxy for uncertainty, as least during stable periods. Lahiri and Sheng (2010) give a contrary argument that dispersion is a reliable measure for uncertainty in a stable period but less useful as a proxy in periods with large volatility. Our results highlight the different tendency in dispersion and forecaster uncertainty and illustrate why the studies come up with very different conclusions. Co-movements are observed during crisis period where dispersion and uncertainty both increase significantly, but the persistency of this effect on forecaster uncertainty makes the correlation at long run remain unclear. This explains the different correlation pattern in different economic environment found in the literature.

Finally, this paper contributes to the extant literature that explore factors to explain the forecast bias made by the professionals. Existing studies indicate that analysts' forecasts of corporate earnings do not meet the classical standards of the rational expectations hypothesis and the extent of bias is predictable from publicly available information (e.g., Ackert and Athanassakos, 1997; Das, Levine, and Sivaramakrishnan,1998), and is greater under higher degree of information uncertainty (e.g. Zhang, 2010). In addition, the moods or psychological states of people have been shown to affect their judgment and professional forecasts are no exception (e.g. Dolvin et al., 2009; Lo and Wu, 2010). For example, Dong and Heo (2014) provide evidence of distraction effect by showing that a higher flu intensity is associated with a lower degree of disagreement among analysts, while pessimism, one of the mood proxy variables, results in lower analyst optimism and leads to more pessimistic forecasts among analysts in New Orleans than among analysts outside of Louisiana following Hurricane Katrina in 2005 (Bourveau and Law, 2016). Our results suggest that the existence of behavioral bias can be extended to the context of professional economists by providing evidence of framing effect in surveys of professional forecasters. This effect has important normative implications on survey designers who organize the crucial forecast data and analysist as well as policy makers who would draw reference from the forecasts.

The rest of the paper is organized as follows. Section 3.2 describes the surveys and measures of dispersion and uncertainty examined in this study. Section 3.3 illustrates the features of the survey design and discusses how they change over time in different circumstances. Section 3.4 focus on the main results that how dispersion and uncertainty respond to survey changes. Section 3.5 concludes.

3.2 Data and Measure

3.2.1 Data

The focus of this paper is to explore the relationship between survey design and forecasters' behavior. Table 3.1 summarizes the surveys included in the sample of this paper.

A US-SPF

The earliest and best-known density forecasts are from the Survey of Professional Forecasters (US-SPF), a quarterly survey formerly initiated by the American Statistical Association (ASA) and the National Bureau of Economic Research (NBER) in 1968Q4 and took over by the Federal Reserve Bank of Philadelphia in 1992Q2. The ranges of the survey questions for density forecasts and the definition of the variables to predict changed several times. First, we only consider the forecasts starting with the 1992Q1 survey, even though the point and density forecasts for output and inflation are available all the way back to the first survey in 1968Q4. There are several reasons for using the post-1992Q1 sample period. First, there were several definition changes to the variables forecasted prior to the 1992Q1 survey (for output, from nominal GNP to real GNP to real GDP). In some years prior to 1992Q1 the survey covered forecasts for the previous and current year, instead of the current and following year. Since

1992Q1 the definitions have been stable. Second, in the surveys during 1980s (1981Q3 to 1991Q4), the interval widths given for density forecasts switched from one percentage point to two percentage point, leading to much cruder density forecasts. In addition to this, the intervals provided in some of the earlier surveys were sometimes completely misaligned with the expectations of the forecasters, resulting in density forecasts with probabilities concentrated in the first or last open-ended bins. In contrast, the sample period selected in this paper is much cleaner in terms of variable and forecast definitions, and has fewer instances of ‘misaligned’ bins. The width of interior intervals for GDP growth is 1 percent while for inflation, it changed from 1 percent to 0.5 percent in 2004Q1. The horizon structure for SPF is quite straightforward and consistently includes only fixed event, where respondents are required to provide their expectations for current year and following year. Though the survey extended the annual forecast horizon two years for real GDP growth in 2009Q2, only current year and following year forecasts are considered in this analysis as they consistently showed up during the whole sample period (fixed-event scheme):

1. calendar year (horizon changes from 3 to 0 from Q1 to Q4),
2. next calendar year (horizon changes from 7 to 4 from Q1 to Q4),

B ECB-SPF

Since 1999 the European Central Bank (ECB) has conducted a quarterly survey of expectations for some of the euro-area key macroeconomic variables. The survey is known as the ECB’s Survey of Professional Forecasters (ECB-SPF). In this survey, inflation is defined as the year-to-year percentage change of the Harmonised Index of Consumer Prices (HICP) published by Eurostat and real gross domestic product growth is defined as the year-to-year percentage change of real GDP, based on standardised ESA definition. The horizon structure of ECB-SPF is more complicated than that of US-SPF, containing both fixed-event and fixed-horizon scheme. Since the horizon structure changes over time, we focus on the following that are consistently appearing in the sample:

1. current calendar year (horizon changes from 3 to 0 from Q1 to Q4),
2. next calendar year (horizon changes from 7 to 4 from Q1 to Q4),
3. "rolling horizon" for the month (for HICP inflation) and quarter (for GDP growth) one year ahead of the latest available observation at the time of the survey (horizon = 3 for HICP inflation, horizon = 2 for GDP growth),
4. "rolling horizon" for the month (for HICP inflation) and quarter (for GDP growth) two year ahead of the latest available observation at the time of the survey (horizon = 7 for HICP inflation, horizon = 6 for GDP growth).

The "rolling horizon" scheme, which is also a fixed-horizon scheme, is a little tricky in this survey. The horizons are set 1- and 2- years ahead of the period (month or quarter) for which the latest official release of the underlying indicators is available, and therefore could not be simply identified as 4 or 8 but differ across indicators. Take the survey conducted in 2007Q1 for example. The latest release of inflation was for December 2006 so the participants were asked for the year-on-year inflation predictions in December 2007 and December 2008, which corresponds to the horizon of 3 and 7, respectively. For GDP growth, on the other hand, the latest available release related to 2006Q3 and forecasters were asked for expectations for GDP growth in 2007Q3 and 2008Q3, with the horizons being 2 and 6 instead. In short, for the fixed-horizon scheme used in this study, the horizon is 3 and 7 for inflation and 2 and 6 for GDP growth.

As for the details of the range specification for the density forecasts, ECB-SPF has a very flexible adjustment scheme with ranges for density forecasts changing over time. While the width of the interior intervals remains constant at 0.5 percent for both output growth and inflation.

C SEF

Dating back to 1996Q2, the Bank of England initiated a quarterly survey of macroeconomic expectations, known as the Survey of External Forecasters (SEF). The first appearance of this data was in February 1993 where the Inflation Report included information on the point forecasts of inflation made by outside forecasters. From February 1996, in addition to the central projections for inflation, the forecasters are also asked about the probabilities they attach to various possible inflation outcomes. In line with the definition of official inflation target, inflation in the survey was first defined as the Retail Prices excluding mortgage interest payments (RPIX), and switched to the Consumer Prices Index (CPI) from the 2004Q1 survey, following the change in the Bank's official target measure in December 2003. As for GDP growth, the point and density forecasts have appeared since 1998Q1.

There is a break in the horizon structure of SEF 2006Q2. Before 2006Q2, it asked the forecasters about their predictions at three future points in time (combining fixed-event (1,2) and fixed-horizon (3) schemes):

1. the fourth quarter of the current year (horizon changes from 3 to 0 from Q1 to Q4);

2. the fourth quarter of the following year (horizon changes from 7 to 4 from Q1 to Q4);
3. the corresponding quarter two years ahead (horizon = 8).

While since 2006Q2, the survey changed to pure fixed-horizon scheme:

1. the corresponding quarter one year ahead (horizon = 4);
2. the corresponding quarter two years ahead (horizon = 8);
3. the corresponding quarter three years ahead (horizon = 12).

Similar to the other two datasets, the survey design of density forecasts, in terms of the number of bins, the upper bound and lower bound, and the width of the intervals, is evolving over time. The details of the change scheme will be presented in the next section.

3.2.2 Sample for Analysis

All of the three datasets assign a unique identification number to each respondent, so that their individual responses can be tracked over time and their point and density forecasts can be matched. This enables us to limit our sample so that we only include density forecasts where the median of a forecaster's the density forecast matches the forecaster's point forecasts. The matching method we use is based on the bounds implied by the density forecasts (Engelberg, Manski and Williams, 2009). The first step to construct the matched sample is to calculate the lower and upper bounds of both subjective median and mean. The interval where the median lies in can be obtained from the probabilistic responses directly. To calculate lower and upper bounds on the subjective mean, we assume that each bin's probability mass is placed at the bin's lower and upper endpoint respectively. The results are then generated by averaging the lower and upper endpoints weighted by the probabilities. If the point forecast is located within any of the two sets of bounds, the density forecast is counted as ‘matched’ to the point forecast. The matching takes care of another issue in the sample, that is the presence of outliers and unusual observations that appear to be errors of some sort, or are at least difficult to otherwise justify. These outliers are also removed via the matching process.

3.2.3 Measures

As in the previous chapter, we use a set of statistics to describe the characteristics of a density forecast. We use $f_i(Y_{t,h})$ to denote forecaster i ’s period t density forecast of annual GDP growth or annual inflation made h-quarters ahead, $h = 0,1,...,7$. The subscript t is a quarterly date index (1992Q1, 1992Q2, etc.) and represents the survey date. As examples, $f_i(Y_{1992Q4,0})$ is forecaster i ’s estimate of annual GDP Growth or Inflation over year 1992, made in the last quarter of 1992, whereas $f_i(Y_{1992Q4,4})$ denotes the forecaster’s prediction for year 1993. The target year is not represented in this notation, and must be derived from the survey date t and the horizon. We use the sample starting $t = 1992q1,...,2016q4$, where $t-1$ refers to the previous quarterly date. We summarize each density forecast using measures of location and spread. One feature of the density forecasts is that the end bins are open-ended, which complicates the computation of moment-based and entropy-based statistics. In this paper, we use instead the median $M_{i,t,h}$, the central 90% range $R_{i,t,h}$:

$$M_{i,t,h} = x_{50,i,t,h}$$

$$R_{i,t,h} = x_{95,i,t,h} - x_{5,i,t,h}$$

where x_α represents the α-th percentile. Here we use instead the median and the central 90% range, whereas the range might be considered a measure of individual uncertainty.

Our main reason for using percentile-based descriptions of the density forecasts is to avoid the strong assumptions that are needed for moment-based and entropy-based statistics. Of course, the percentile-based statistics that we use also has its disadvantages, e.g., it requires interpolation within the bins (we use linear interpolation). In addition, the 5th (95th) percentiles cannot be computed if the probabilities reported for the first or last bins are greater than 5 (probabilities are reported out of 100). This interpolation is also an assumption regarding the shape of the distribution, though perhaps a less critical one, as the assumption is only required in the bins in which the 5th, 50th, and 95th percentiles falls. The fact that the 5th and 95th percentiles cannot be computed in some cases is also not a major issue for our sample period. Finally, recent papers have considered moment- and entropy based statistics (Rich and Tracy 2010, Boero et. al. 2008, 2015) for some of the issues we examine, so it is interesting to see how our results compare when using different measures.

3.3 Survey Design

In this section, we focus on the changing scheme of survey design of the datasets. Figure 3.1 plots the ranges of interior intervals together with the fluctuations of corresponding values of underlying indicators. Several important features are highlighted in these figures. First, in addition to changes to survey design that are necessitated by inherent changes in the variables predicted, we observe some adjustments of forecast bins that are not prompted by any structural change of the underlying indicators. For example, the SPF designer reduced both the whole range and interval width of the bins provided in 2014Q1 without any significant variation in inflation. This set of changes is neutral to the economic environment and occurs when the current ranges are consistently larger than the variation of the real indicators during a period and the designers shrink the total range, sometimes together with the width of the intervals, to better fit the possible amplitude of the indicator's movement. This category of changes provides a clean setting that allows us to distinguish the framing effect from the effect of real economic fluctuation. Second, even in response to the economic environment, the surveys are quite different in the adjustment strategies. Among the three surveys, ECB-SPF are more sensitive to the changes in the underlying indicators while SEF is most reluctant to do modifications. For example, during 2008-2009, all the countries and regions in the sample underwent a big drop in output growth with the real value falling far beyond the lower bound of the interior intervals in all three surveys. After this surprising shock, ECB-SPF adjusted the ranges twice to completely include the full range of indictor's values in the inner intervals. However, SPF and SEF only slightly adjusted the intervals by 0.01 to 0.02.

The details of the variations in survey design are depicted in Figure 3.2. For US-SPF, there is only one change in the design scheme within sample period for both inflation and GDP growth. The changing scheme in GDP growth survey is driven by the crisis when there is a sudden drop in real GDP growth while the change in inflation survey question is most likely a modification of the survey design to best fit the volatility of the underlying variable.

1. Inflation: [<0,0-1,1-2,2-3,3-4,4-5,5-6,6-7,7-8,>8] changed to [<0,0-0.5,0.5-1,1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,3.5-4,>4] in 2014Q1;

2. GDP growth: [<(-2),(-2)-(-1),(-1)-0,0-1,1-2,2-3,3-4,4-5,5-6,>6] changed to [<(-3),(-3)-(-2),(-2)-(-1),(-1)-0,0-1,1-2,2-3,3-4,4-5,5-6,>6] in 2009Q2;

ECB-SPF, just like what we observed in Figure 3.1(b), is always ready to adjust the bounds of interior intervals to fully cover the real variation of the variable predicted. As a result, there are several changes in our sample period and all of the adjustments are related to the real economy volatility.

1. Inflation:
    a. From 1999Q1: [<0,0-0.5,0.5-1, 1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,>3.5]
    b. From 2000Q4: [<0,0-0.5,0.5-1, 1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,3.5-4,>4]
    c. From 2001Q1: [<0,0-0.5,0.5-1, 1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,>3.5]
    d. From 2008Q3: [<0,0-0.5,0.5-1, 1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,3.5-4,>4]
    e. From 2009Q2: [<(-2),(-2)-(-1.5),(-1.5)-(-1),(-1)-(-0.5),(-0.5)-0,0-0.5,0.5-1, 1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,>3.5]
    f. From 2010Q1: [<(-1),(-1)-(-0.5),(-0.5)-0,0-0.5,0.5-1, 1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,>3.5]
2. GDP growth:
    a. From 1999Q1: [<0,0-0.5,0.5-1,1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,3.5-4,>4]
    b. From 2000Q2: [<0,0-0.5,0.5-1,1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,3.5-4,4-4.5,4.5-5,>5]

    c. From 2001Q1: [<0,0-0.5,0.5-1,1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,3.5-4,>4]
    d. From 2008Q4: [<(-1),(-1)-(-0.5),(-0.5)-0,0-0.5,0.5-1,1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,3.5-4,>4]
    e. From 2009Q2: [<(-6),(-6)-(-5.5),(-5.5)-(-5),(-5)-(-4.5),(-4.5)-(-4),(-4)-(-3.5),(-3.5)-(-3),(-3)-(-2.5),(-2.5)-(-2),(-2)-(-1.5),(-1.5)-(-1),(-1)-(-0.5),(-0.5)-0,0-0.5,0.5-1,1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,3.5-4,>4]
    f. From 2010Q1: [<(-1),(-1)-(-0.5),(-0.5)-0,0-0.5,0.5-1,1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,3.5-4,>4]

The change pattern is rarer in SEF survey, where the changes in the early (pre-crisis) period are not relevant to fluctuations of underlying variables and the adjustments in the crisis period are mostly made in response to the changing of the economic environment.

1. Inflation:
    a. From 1996Q2: [<1,1-2.5,2.5-4,4-5.5,>5.5]
    b. From 1998Q1: [<1.5,1.5-2.5,2.5-3.5,>3.5]
    c. From 1999Q2: [<1.5,1.5-2,2.5-3,3-3.5,>3.5]
    d. From 2004Q1: [<1,1-1.5,1.5-2,2-2.5,2.5-3,>3]
    e. From 2009Q1: [<0,0-1, 1-1.5,1.5-2,2-2.5,2.5-3,>3]
2. GDP growth:
    a. From 1998Q1: [<0,0-1,1-2,2-3,3-4,>4]
    b. From 2004Q1: [<1,1-2,2-3,>3]
    c. From 2008Q4: [<0,0-1,1-2,2-3,>3]
    d. From 2009Q1: [<(-1),(-1)-0,0-1,1-2,2-3,>3]

3.4 Main Results

The focus of this section is how professional forecasters' behavior changes in response to the adjustments of survey design. We examine two important measures that attract close attention of the economists: dispersion and forecaster uncertainty. Dispersion is defined as the standard error of the medians of the forecasters' density forecasts:

$$D_{t,h} = std.dev.(M_{i,t,h})$$

and forecaster uncertainty is defined as the mean of the forecasters' 90% range, a proxy of the individual forecaster uncertainty:

$$U_{t,h} = mean(R_{i,t,h})$$

A US-SPF

Figure 3.3 shows an overview of how the dispersion and uncertainty vary along the survey date. The vertical line in each plot indicates a change in the survey design. The subplots in the left column relates to the dispersion of medians of density forecasts and those in the right column relates to the average forecaster uncertainty reported by the density forecasts. The top two rows depict the measures of inflation forecasts for current year and following year respectively and the bottom two show the measures of output growth.

From the top four figures drawing features of inflation forecasts, we observe only one change of the survey in 2014Q1. The change is not driven by the inherent change in inflation with the coverage shrinking from 0-8% to 0-4% and interval width decreasing from 1% to 0%. There are significant differences in the response of dispersion and inflation to this change. Dispersion does not show any change in the moving pattern while uncertainty experiences a sudden and persistent drop at the change point when the surveyor shrinks both coverage and width of the bins provided to the forecasters. This finding is of great importance as this change is not correlated with any economic variable and thus the response serves as a strong evidence of behavioral bias.

As for the GDP growth, the only survey change occurred in 2009Q2, when the lower bound of the bins decreased slightly from -2 to -3. This change is triggered by the crisis where the real output growth fell suddenly beyond the original lower bound -2. Dispersion increases immediately after the crisis being observed and the survey design being modified and returns to normal level after the crisis. At the same time, uncertainty also increases slightly after the crisis.

The two survey changes are corresponding to two different cases as we defined earlier: the one in inflation survey is independent of underlying variable while the one in output survey is responding to the real economic variation. While it is intuitive that dispersion and uncertainty increase during the crisis, the unusual and significant rise in forecaster uncertainty of inflation at 2014Q1 provides strong evidence to framing effect that the survey scheme itself affects the forecasters' behavior.

B ECB-SPF

Compared to the other two surveys, ECB-SPF is more sensitive to the changes in the underlying indicators and all the adjustments are made in line with the real economy. This provides us with a direct comparison of how dispersion and uncertainty respond to the survey change. An overview of the two measures is depicted in Figure 3.4. There are four forecasts at each survey date included in our sample, corresponding to the four rows of the plots: forecast for the current calendar year; the next calendar year; the month (for HICP inflation) and quarter (for GDP growth) one year ahead of the latest available observation at the time of the survey; the month (for HICP inflation) and quarter (for GDP growth) two year ahead of the latest available observation at the time of the survey. The left two columns show dispersion and uncertainty of inflation and the right two columns show dispersion and uncertainty of output growth. As the survey changes are all in line with the economic variations, there is no significant difference for the two variables of inflation and output. Both dispersion and uncertainty respond to real economic fluctuations and the measures increase as the economic variables become more volatile. However, the dispersion usually goes back to the normal level soon after the shock and is more volatile while changes in uncertainty are more persistent. These findings are consistent with what we have shown in Chapter two that dispersion is highly correlated with macroeconomic uncertainty while forecaster uncertainty are more subjective and less linked with macroeconomic uncertainty.

The fixed-horizon scheme exploited in the ECB-SPF allows us to conduct the fixed-horizon analysis on the relationship between the measures and horizons. Figure 3.5 provides a clear comparison of the features of dispersion and uncertainty, with measures of inflation in the first row and measures of output growth in the second row. The dispersion is depicted in the left column and uncertainty is depicted in the right column. Each graph plots two different horizons (3 and 7 for inflation and 2 and 6 for output growth). Both dispersion and uncertainty respond to real economic change while effect in uncertainty is more persistent. Another interesting fact highlighted in this figure is that disagreement in long horizons is usually smaller than that at short horizons while uncertainty is totally opposite: long-horizon uncertainty is always larger than short-horizon one. This indicates that at long horizon where information is quite limited, the forecasters are more uncertain but disagree less in the first moment of expectations.

C SEF

The survey scheme in SEF changed several times within our sample period and some changes, especially in the pre-crisis period appear to be independent of the underlying indicators while the changes in the crisis period are driven by the real economic volatility. This feature provides another clean setting that allows us to differentiate the effect of structural break in real economy from the behavioral effect.

Figure 3.6 is an overview of the responses. The changes in the survey are highlighted by the vertical lines. There are three forecasts at each survey date in SEF but the horizon scheme changed in 2006Q2 from combined fixed-event and fixed horizon scheme to pure fixed-event scheme. So we are left with five different series in the sample, which are illustrated in the five rows from the top to the bottom: forecast for the fourth quarter of the current year (before 2006Q1); the fourth quarter of the following year (before 2006Q1); the corresponding quarter one year ahead (after 2006Q2); the corresponding quarter two years ahead (full sample); the corresponding quarter three years ahead (after 2006Q2). The left two columns show dispersion and uncertainty of inflation and the right two columns show dispersion and uncertainty of output growth. As we noted in the previous section, the two changes in 1998Q1 and 1999Q2 are not responding to any structural break in the economy and completely independent of the values of inflation. There is no clear response in dispersion as shown in the first column, however, both changes lead to steep decline in forecaster uncertainty, illustrated in the second column. As for the changes in the crisis that are forced by the economic environment, it is not surprising that both dispersion and uncertainty rise immediately.

Special attention must be paid to GDP growth on a quite abnormal change in 2004Q1, when the economy is quite volatile in the following period but the survey designer changed the bin setting from [<0,0-1,1-2,2-3,3-4,>4] to [<1,1-2,2-3,>3]. Opposite to the intuition that more bins and wider coverage should be provided in volatile period to better capture the views of forecasters about future economy, the survey reduced both coverage and number of bins. The response is also very interesting: the dispersion of medians of density forecasts declines and forecaster uncertainty goes up. Though the results might be affected by the limited sample (we lose some observations after matching and drop some records without valid medians due to the limitation of the survey design), the response is enough to provide a note of caution to the data designer and user: poorly designed surveys severely limit the performance of forecasts.

We also conduct the fixed-horizon analysis for SEF. Figure 3.7 compares the features of dispersion and uncertainty, with measures of inflation in the first row and measures of output growth in the second row. The dispersion is depicted in the left column and uncertainty is depicted in the right column. Each graph plots three different horizons (4, 8, and 12), with horizon 8 showing up in the data for full period. Again, both dispersion and uncertainty respond to real economic change while effect in uncertainty is more persistent. Dispersion in short horizon is usually larger than that for long horizon especially in the unstable period while uncertainty is completely opposite. The seemingly unreasonable change in 2004Q1 leads to lower dispersion and higher uncertainty during that period, suggesting a worse performance of forecasts restricted by the survey design.

In summary, the results are consistent with what we observe in the US-SPF and ECB-SPF datasets, in that both disagreement and uncertainty increase when there is a shock to the macroeconomic variables, while the effect on disagreement is more transient, and on uncertainty more persistent. Dispersion can be affected by real economic change, regardless of the bin changes, but does not move with the survey changes that are neutral to economic environment. Uncertainty responds to both cases of survey adjustment and the effect will last for a long period after the shock. The output growth survey in SEF provides a rare survey change case where the modification goes against the evolution of real economy in 2004Q1, and our analyses suggest that improper design could worsen the performance of density forecasts.

3.5 Concluding Remarks

In this paper, we examine how the survey design influences the forecasters' behavior. With a comprehensive dataset including three main surveys worldwide, namely US-SPF, ECB-SPF, and SEF, we are able to detect and analyze changes of the density forecast bins. While the adjustment of forecast bins reasonably arises from the fluctuation of underlying macroeconomic variable, there are also cases where the modification is neutral to the economic environment. The inclusion of three datasets allows us to distinguish response patterns of professional forecasters under different categories of survey adjustments. The results suggest that disagreement only responds to changes caused by real economy. Uncertainty responds to both categories of survey changes and the effect is more persistent. These empirical facts highlight the importance of behavioral perspective when inferences are drawn from professional forecasters (Clark, Ghosh and Hanes, 2018). A more useful study of this issue is essential if we are to draw more accurate inferences from such surveys especially regarding forecast uncertainty.

Table 3.1 A Summary on Surveys of Professional Forecasters.

	US-SPF	ECB-SPF	SEF
Survey organization	Orginally ASA/NBER; currently the Pheladelphia	European Central Bank	Bank of England
Average number of respondents in matched sample	30 (after 1992Q1)	52	25
Starting date	1968Q4 sample from 1992Q1	1999Q1	1996Q2 for inflation 1998Q1 for output growth
Periodicity	Quarterly	Quarterly	Quarterly
Horizon structure	Fixed target	Fixed target and fixed horizon	1996Q2-2006Q1 Fixed target and fixed horizon 2006Q2-present Fixed horizon
Inflation expectation	GDP deflator	Euro area HICP	RPIX before 2003Q4 CPI after 2004Q1

Notes: A summary of the surveys included in the sample. Three surveys are included: US-SPF, ECB-SPF and SEF.

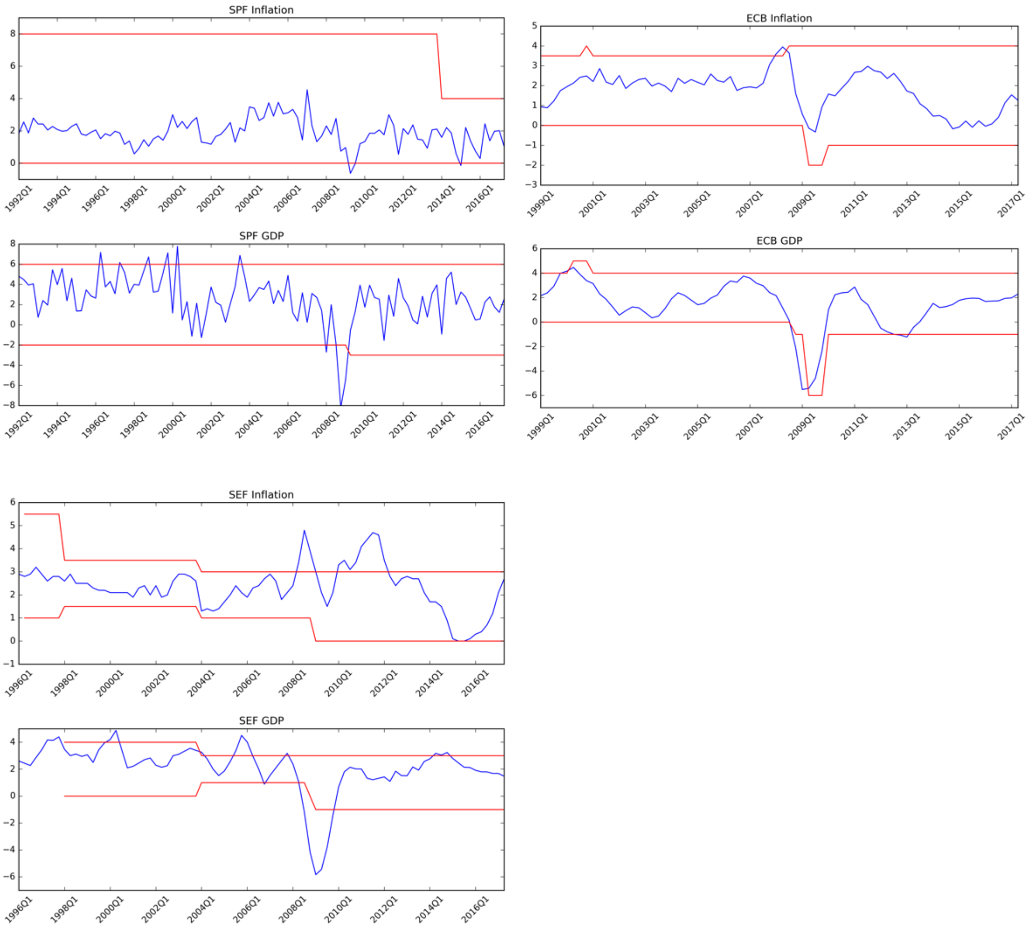

Figure 3.1 Interior Bounds and Fluctuation of Underlying Indicators

Ranges of interior intervals and the fluctuations of corresponding values of underlying indicators. The two horizontal lines indicate the interior bounds of the bins provided by the survey to the forecasters and the curve shows the series of the underlying variables being predicted. The left subfigure in the top row relates to US-SPF and the right one relates to ECB-SPF. The one in the bottom relates to SEF.

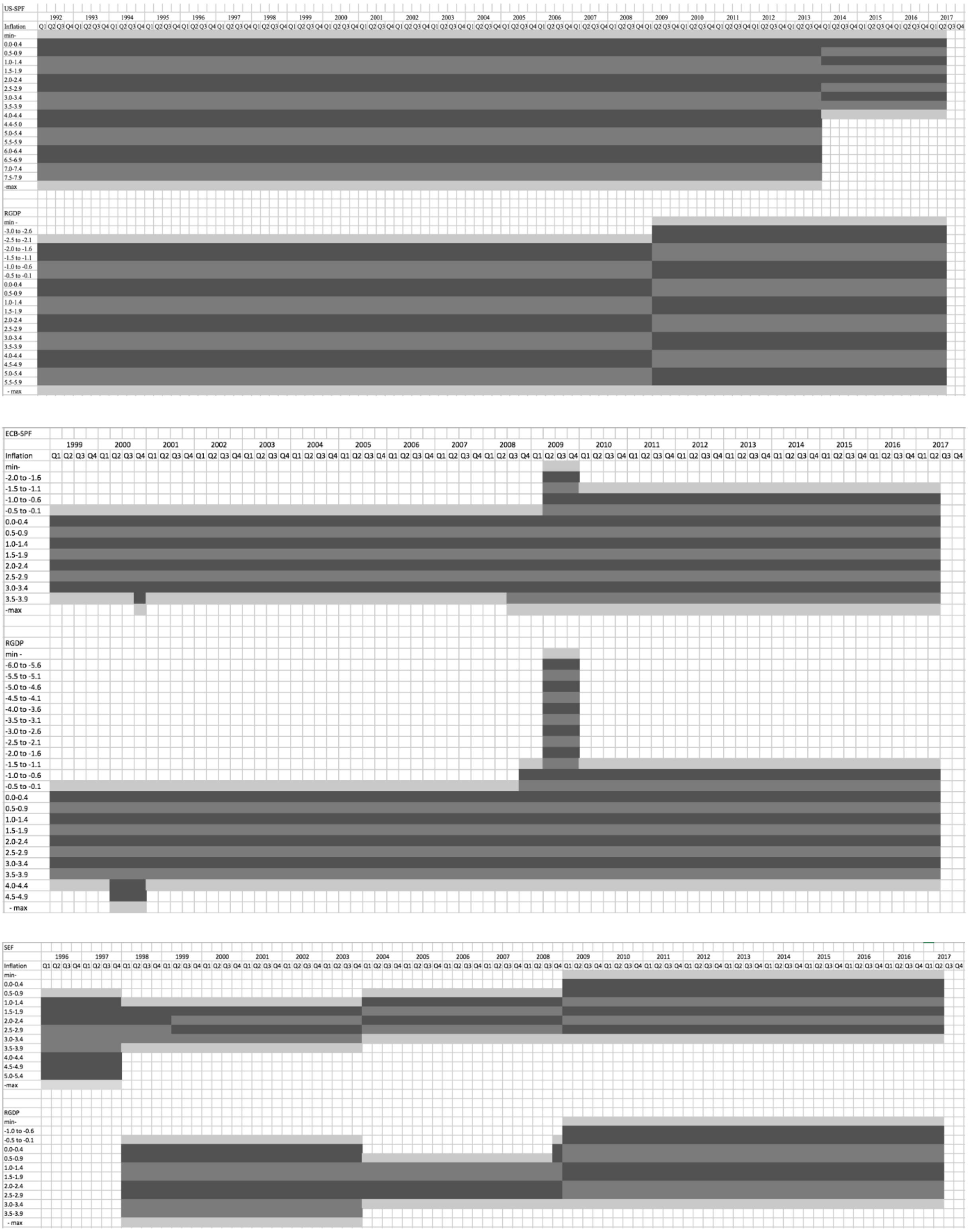

Figure 3.2 Survey Design Variations

The details of the variations in survey design, with the first one corresponds to the details of US-SPF, the second one to ECB-SPF and the last one to SEF. In each subfigure, the top part shows variations of inflation survey and the bottom part shows that of output survey. The different color highlights the width of the interior bins.

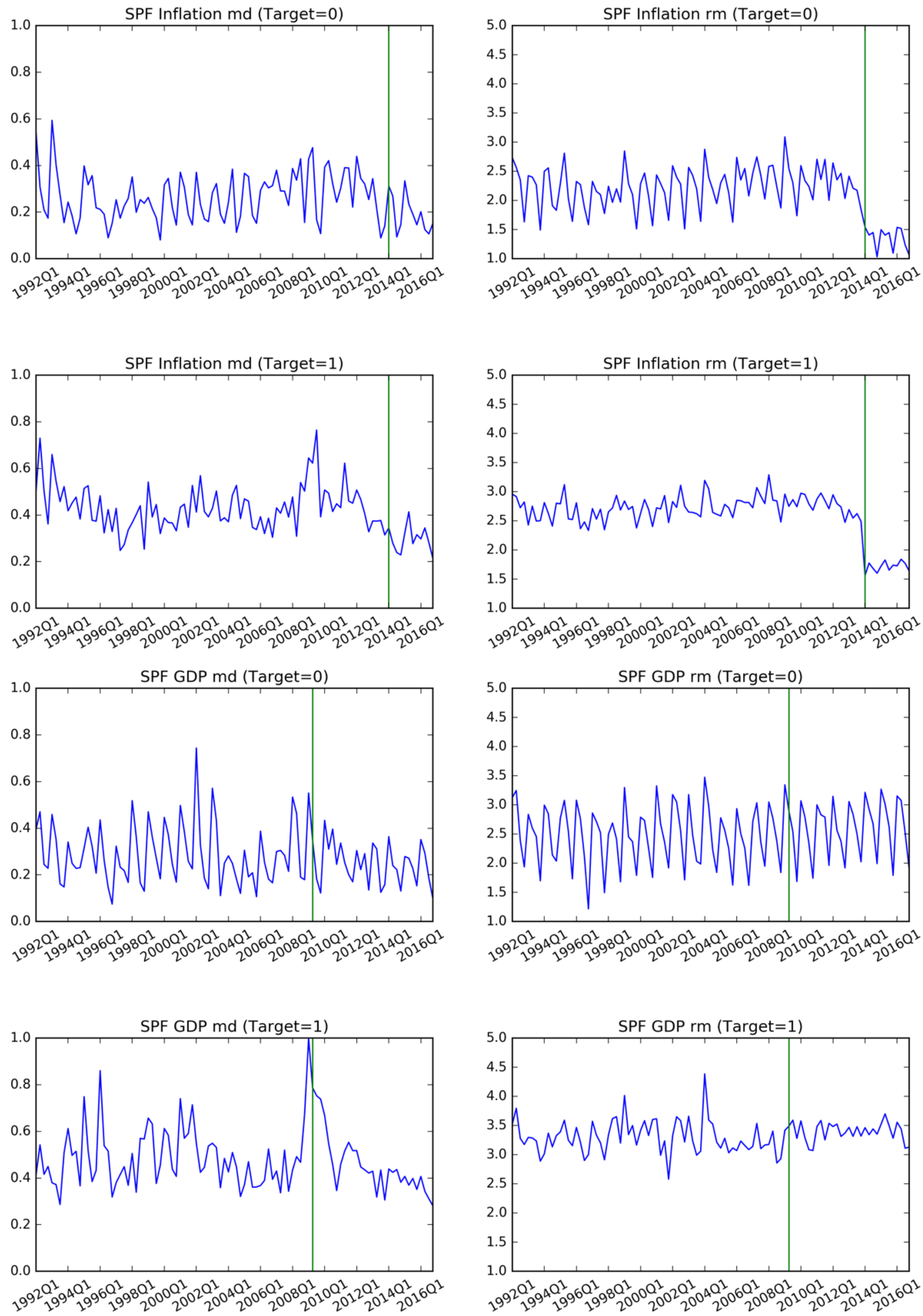

Figure 3.3 An overview of US-SPF

An overview of how the dispersion and uncertainty vary along the survey date in US-SPF. The vertical line in each plot indicates a change in the survey design. The subplots in the left column relates to the dispersion of medians of density forecasts and those in the right column relates to the average forecaster uncertainty reported by the density forecasts. The top two rows depict the measures of inflation forecasts for current year and following year respectively and the bottom two show the measures of output growth.

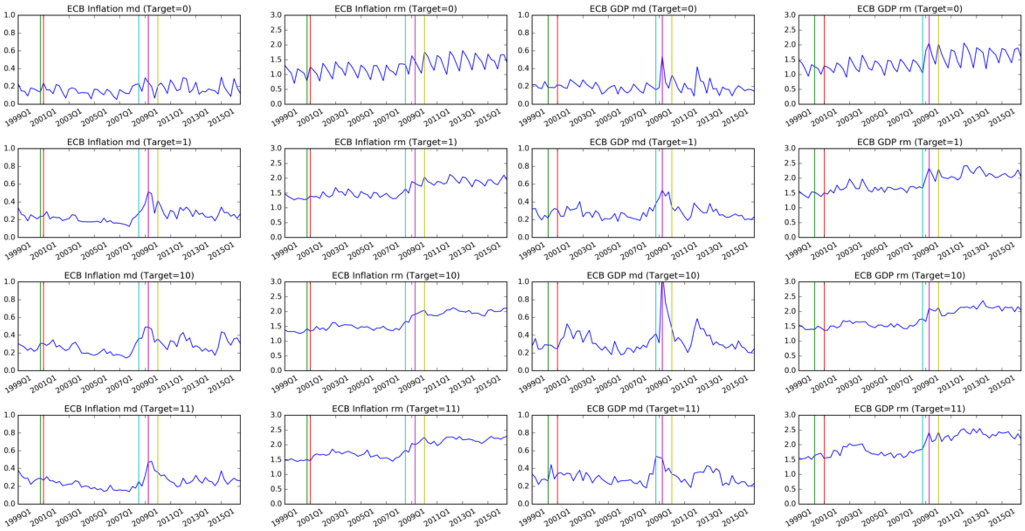

Figure 3.4 An Overview of ECB-SPF

An overview of how the dispersion and uncertainty vary along the survey date in ECB-SPF. The vertical line in each plot indicates a change in the survey design. The four rows correspond to the four forecasts at each survey date: forecast for the current calendar year; the next calendar year; the month (for HICP inflation) and quarter (for GDP growth) one year ahead of the latest available observation at the time of the survey; the month (for HICP inflation) and quarter (for GDP growth) two year ahead of the latest available observation at the time of the survey. The left two columns show dispersion and uncertainty of inflation and the right two columns show dispersion and uncertainty of output growth.

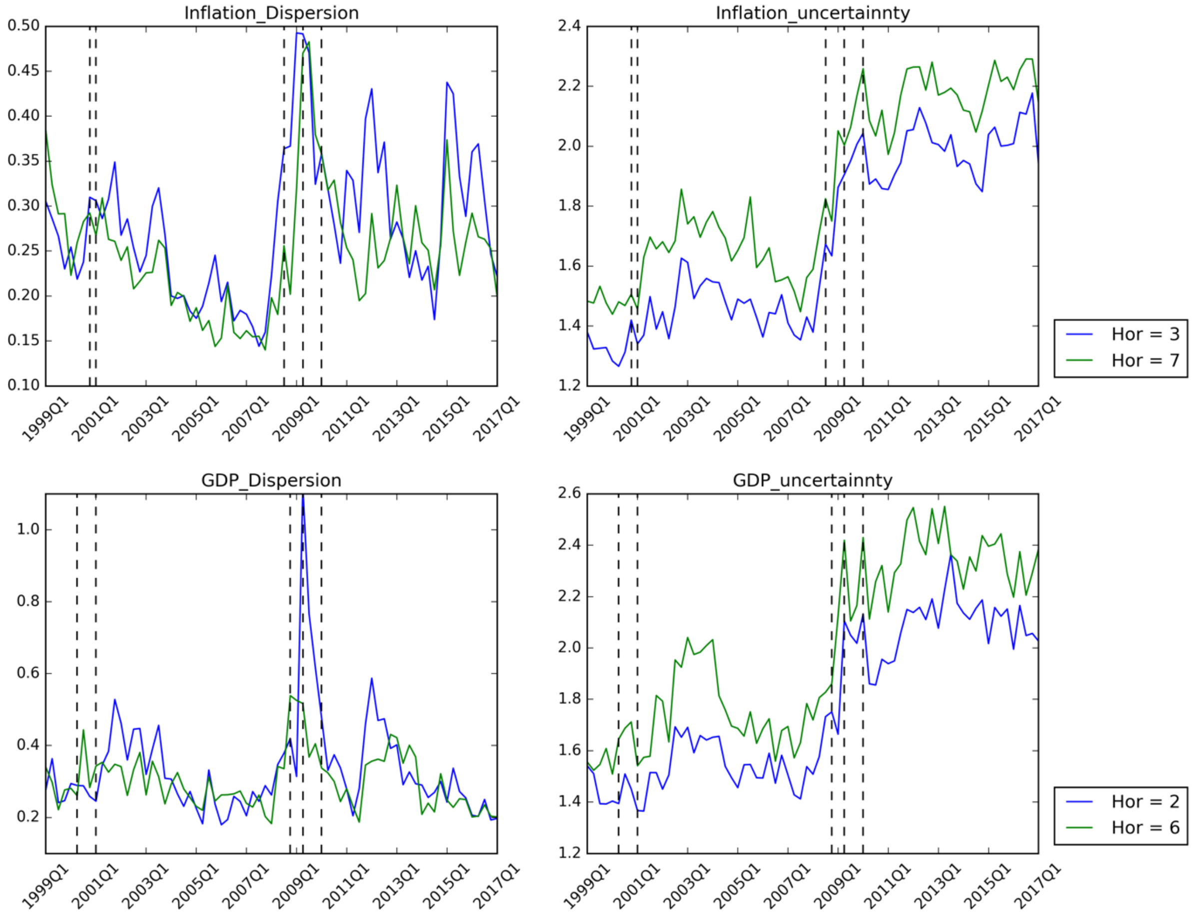

Figure 3.5 ECB-SPF: Fixed-horizon Analysis

Comparison of the features of dispersion and uncertainty, with measures of inflation in the first row and measures of output growth in the second row. The dispersion is depicted in the left column and uncertainty is depicted in the right column. Each graph plots two different horizons (3 and 7 for inflation and 2 and 6 for output growth).

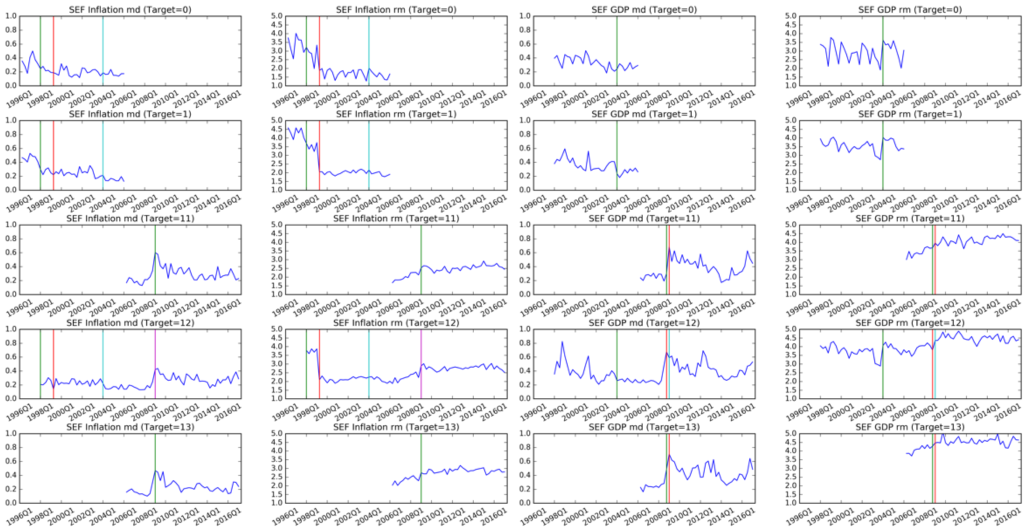

Figure 3.6 An Overview of SEF

An overview of how the dispersion and uncertainty vary along the survey date in SEF. The vertical line in each plot indicates a change in the survey design. The five rows from top to bottom correspond to the five different series of forecasts in the sample: forecast for the fourth quarter of the current year (before 2006Q1); the fourth quarter of the following year (before 2006Q1); the corresponding quarter one year ahead (after 2006Q2); the corresponding quarter two years ahead (full sample); the corresponding quarter three years ahead (after 2006Q2). The left two columns show dispersion and uncertainty of inflation and the right two columns show dispersion and uncertainty of output growth.

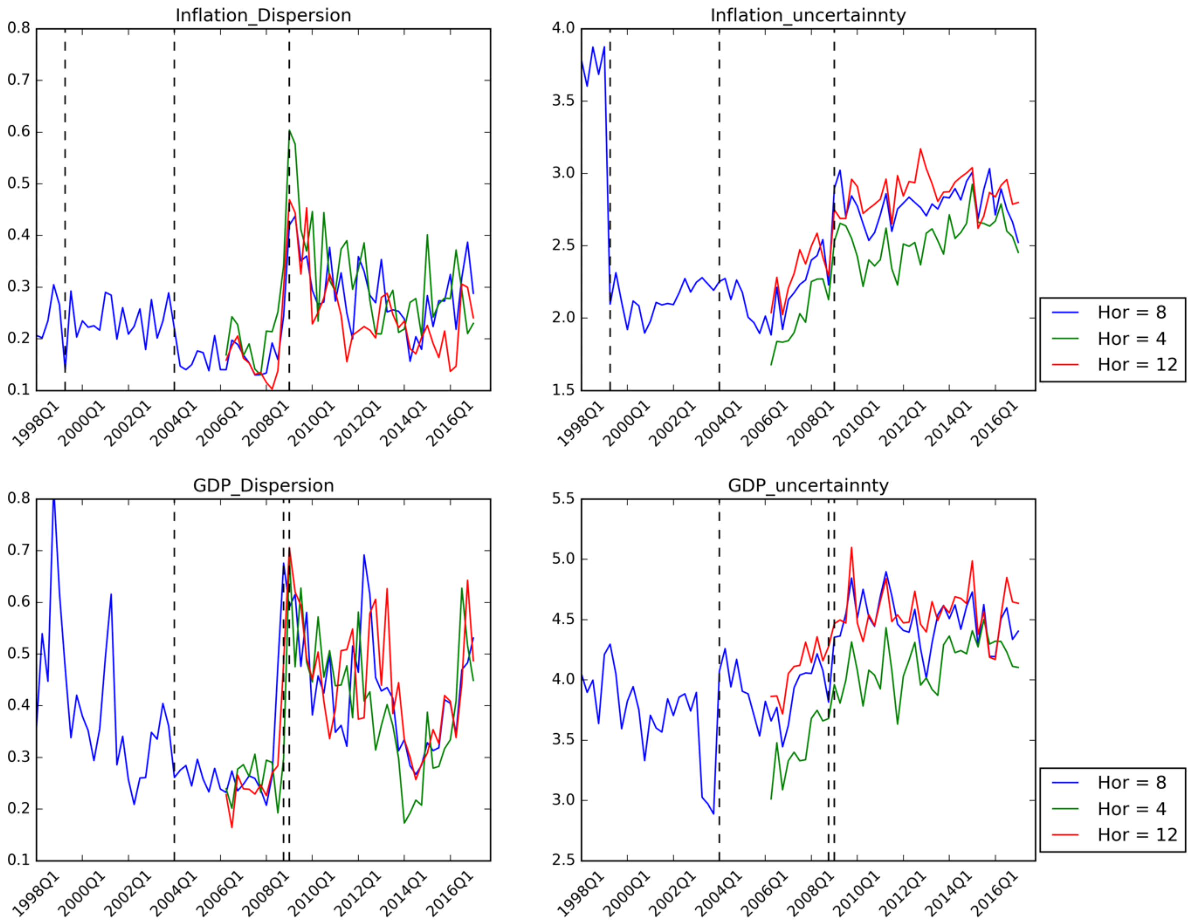

Figure 3.7 SEF: Fixed-horizon Analysis

Comparison of the features of dispersion and uncertainty, with measures of inflation in the first row and measures of output growth in the second row. The dispersion is depicted in the left column and uncertainty is depicted in the right column. Each graph plots three different horizons (4, 8, and 12), with horizon 8 showing up in the data for full period.

Chapter 4 Concluding Remarks and Future Research

4.1 Concluding Remarks

This dissertation focuses on the two key features of density forecasts, dispersion and uncertainty, and explore underlying determinants of the features using three world-wide surveys, US-SPF, ECB-SPF and SEF. It contributes to the literature summarized in Chapter One through additional dimensions in terms of the effect of uncertainty in dispersion and behavioral perspective.

Chapter two explores the role of uncertainty in explaining dispersion in professional forecasters' density forecasts of real output growth and inflation. We consider three separate notions of uncertainty: general macroeconomic uncertainty (the fact that macroeconomic variables are easier to forecast at some times than at others), policy uncertainty, and forecaster uncertainty. We find that dispersion in individual density forecasts is related to overall macroeconomic uncertainty and policy uncertainty, while forecaster uncertainty (which we define as the average in the uncertainty expressed by individual forecasters) appears to have little role in forecast dispersion. On one hand, our work contributes to the literature by extending the analysis on dispersion from point forecasts to density forecasts. Though there is a growing literature focusing on density forecast evaluation and forecaster uncertainty, very few studies examine the dispersion pattern in density forecasts. On the other hand, we exploit various forms of uncertainty to explain forecast dispersion, including both subjective and objective measures showing different perspective of the economic environment. By doing so, we are able to add new elements to the literature examining determinants of dispersion. Furthermore, further analysis on forecaster uncertainty and dispersion in this study highlights different features of these two important measures represented in density forecasts and helps us better understand the relationship of them. In order to test the validity of our results from three dimensions, we conducted several robustness analyses. First, we check the results using sample excluding this period to see whether the key findings are still present in the results and some confusing patterns potentially resulted from the spike in macroeconomic uncertainty would disappear. The results confirm our conclusion that forecaster uncertainty is hardly correlated with either PU or MU. Second, we examine the components of macroeconomic uncertainty, financial uncertainty and real uncertainty, to see which elements really matter in explaining dispersion. The tests reinforce our findings that macro uncertainty plays an important role in determining dispersion of density forecasts. By further examining the components of macro uncertainty, we find that dispersion in medians of inflation forecasts is mainly affected by uncertainty in macro indicators while that of output growth forecasts is more sensitive to uncertainty in financial series. Finally, we add the components of policy uncertainty to check the robustness of our results and identify potential differences in their impact on disagreement. The results suggest that, among all the components, national security and trade policy have little effect in explaining dispersion, while financial regulation has strong effect that dominates MU. The rest of the factors show similar results to our main analysis. The analysis in this part is consistent with our main results that dispersion is strongly correlated to the macro and policy uncertainty.

Chapter three examines the relationship between survey design and forecasters' behavior by exploiting changes to the probability bins supplied to forecasters when soliciting density forecasts. While the adjustment of forecast bins can reasonably arise from the fluctuation of underlying macroeconomic variable, there are also cases where the modification is neutral to the economic environment. Our analysis examines how disagreement and forecaster uncertainty respond to the two different categories of survey changes. The results suggest that disagreement only responds to changes caused by real economy. Uncertainty responds to both and the effect is more persistent. This chapter contributes to the literature on characteristics of forecaster uncertainty from the behavioral perspective. More importantly, it provides a note of caution on the survey design: as the survey design itself could affect the forecasters' behavior, we should pay special attention on the form of the survey questions and be very cautious when making any modifications.

In addition to a summary of the main chapters, this chapter proposes further research plan on investigating the features of forecaster uncertainty using probability integral transforms ("z-statistics"), a commonly-used test for density forecast optimality. The focus is on the shape of the distribution of z-statistics, which is informative about the confidence level as well as bias (optimism or pessimism) of forecasters. There is evidence of significant hysteresis and that survey scheme greatly affects the performance of density forecasts. This study would provide a new application of the traditional density forecast evaluation method and add to the literature an innovative prospective on analyzing the determinants of key features of density forecasts.

4.2 Future Research

This section further examines the features of density forecasts from the perspective of z-statistics, the probability integral transforms. Z-statistic is first proposed and applied to test optimality of density forecasts by Diebold, Gunther and Tay (1998). From then on, various optimality test methods have been proposed. We are not performing the optimality test in this chapter, but use a more informative format of this statistic, the shape and distributions, to illustrate how forecasters' behavior changes in response to the real economic volatility as well as survey changes.

4.2.1 Data and Methodology

Similar to Chapter Three, we include the three surveys of professional forecasters, US-SPF, ECB-SPF and SEF in our analysis. The respondents are professionals from economic agencies and supposed to be relatively homogeneous in terms of ability and motivation. However, the analysis in the previous chapters indicates that it is not the case. The different design scheme of the three surveys allows us to examine the features of this "bounded rationality" in a deeper manner and study the potential determinants of forecast dispersion and uncertainty. Before discussing the results, we first summarized major differences of the three datasets briefly in terms of horizon scheme and question frame.

A US-SPF

The Survey of Professional Forecasters (US-SPF) is the earliest and best-known density forecasts. It's a quarterly survey formerly initiated by the American Statistical Association (ASA) and the National Bureau of Economic Research (NBER) in 1968Q4 and took over by the Federal Reserve Bank of Philadelphia in 1992Q2. The horizon structure for SPF is quite straightforward and consistently includes only fixed event, where respondents are required to provide their expectations for current year and following year. Though the survey extended the annual forecast horizon by two years for real GDP growth starting from 2009Q2, our analysis is limited to only current year and following year forecasts as they can cover the whole sample period (fixed-event scheme):

1. calendar year (horizon changes from 3 to 0 from Q1 to Q4);
2. next calendar year (horizon changes from 7 to 4 from Q1 to Q4).

In our sample, we only consider forecasts starting with the 1992Q1 survey, as the forecasts in this sample period is much cleaner in terms of both variable and forecast definitions. During this period, the specification of questions for density forecasts is very stable and changed only once for GDP growth and inflation respectively.

1. Inflation: [<0,0-1,1-2,2-3,3-4,4-5,5-6,6-7,7-8,>8] changed to [<0,0-0.5,0.5-1,1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,3.5-4,>4] in 2014Q1;
2. GDP growth: [<(-2),(-2)-(-1),(-1)-0,0-1,1-2,2-3,3-4,4-5,5-6,>6] changed to [<(-3),(-3)-(-2),(-2)-(-1),(-1)-0,0-1,1-2,2-3,3-4,4-5,5-6,>6] in 2009Q2.

B ECB-SPF

Since 1999 the European Central Bank (ECB) has conducted a quarterly survey of expectations for some of the euro-zone key macroeconomic variables. The survey is known as the ECB's Survey of Professional Forecasters (ECB-SPF). The horizon structure of ECB-SPF is more complicated than that of US-SPF and contains both fixed-event scheme and fixed-horizon scheme. Considering that the horizon structure changes over time, we focus on the following four of them that are consistently appeared in the sample:

1. current calendar year (horizon changes from 3 to 0 from Q1 to Q4),
2. next calendar year (horizon changes from 7 to 4 from Q1 to Q4),
3. "rolling horizon" for the month (for HICP inflation) and quarter (for GDP growth) one year ahead of the latest available observation at the

time of the survey (horizon = 3 for HICP inflation, horizon = 2 for GDP growth),

4. “rolling horizon” for the month (for HICP inflation) and quarter (for GDP growth) two year ahead of the latest available observation at the time of the survey (horizon = 7 for HICP inflation, horizon = 6 for GDP growth).

The “rolling horizon” scheme, which is also a fixed-horizon scheme, is a little tricky in this survey. The horizons are set 1- and 2- years ahead of the period (month or quarter) for which the latest official release of the underlying indicators is available, and therefore could not be simply identified as 4 or 8 but differ across indicators.

As for the details of the range specification for the density forecasts, ECB-SPF has a very flexible adjustment scheme. The range for density forecasts changes over time while the width of the interior intervals remains constant at 0.5 percent for both output growth and inflation.

1. Inflation:
    a. From 1999Q1: [<0,0-0.5,0.5-1, 1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,>3.5]
    b. From 2000Q4: [<0,0-0.5,0.5-1, 1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,3.5-4,>4]
    c. From 2001Q1: [<0,0-0.5,0.5-1, 1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,>3.5]
    d. From 2008Q3: [<0,0-0.5,0.5-1, 1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,3.5-4,>4]

   e. From 2009Q2: [<(-2),(-2)-(-1.5),(-1.5)-(-1),(-1)-(-0.5),(-0.5)-0,0-0.5,0.5-1, 1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,>3.5]
   f. From 2010Q1: [<(-1),(-1)-(-0.5),(-0.5)-0,0-0.5,0.5-1, 1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,>3.5]

2. GDP growth:
   a. From 1999Q1: [<0,0-0.5,0.5-1,1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,3.5-4,>4]
   b. From 2000Q2: [<0,0-0.5,0.5-1,1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,3.5-4,4-4.5,4.5-5,>5]
   c. From 2001Q1: [<0,0-0.5,0.5-1,1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,3.5-4,>4]
   d. From 2008Q4: [<(-1),(-1)-(-0.5),(-0.5)-0,0-0.5,0.5-1,1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,3.5-4,>4]
   e. From 2009Q2: [<(-6),(-6)-(-5.5),(-5.5)-(-5),(-5)-(-4.5),(-4.5)-(-4),(-4)-(-3.5),(-3.5)-(-3),(-3)-(-2.5),(-2.5)-(-2),(-2)-(-1.5),(-1.5)-(-1),(-1)-(-0.5),(-0.5)-0,0-0.5,0.5-1,1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,3.5-4,>4]
   f. From 2010Q1: [<(-1),(-1)-(-0.5),(-0.5)-0,0-0.5,0.5-1,1-1.5,1.5-2,2-2.5,2.5-3,3-3.5,3.5-4,>4]

C SEF

Initiated by the Bank of England, the first appearance of the Survey of External Forecasters (SEF) was in February 1993 where the Inflation Report included information on the point forecasts of inflation made by outside forecasters. From February 1996, in addition to the central projections for inflation, the forecasters were also asked about the probabilities they attached to various possible inflation outcomes. In line with the definition of official inflation target, inflation in the survey was first defined as the Retail Prices excluding mortgage interest payments (RPIX), and switched to the Consumer Prices Index (CPI) from the 2004Q1 survey, following the change in the Bank's official target measure in December 2003. As for GDP growth, the point and density forecasts have appeared since 1998Q1. There is a break in the horizon structure of SEF 2006Q2. Before 2006Q2, it asked the forecasters about their predictions at three future points in time (combining fixed-event (1,2) and fixed-horizon (3) schemes):

1. the fourth quarter of the current year (horizon changes from 3 to 0 from Q1 to Q4);
2. the fourth quarter of the following year (horizon changes from 7 to 4 from Q1 to Q4);
3. the corresponding quarter two years ahead (horizon = 8).

Since 2006Q2, the survey changed to pure fixed-horizon scheme:

1. the corresponding quarter one year ahead (horizon = 4);
2. the corresponding quarter two years ahead (horizon = 8);
3. the corresponding quarter three years ahead (horizon = 12).

Similar to the other two datasets, the survey design of density forecasts, in terms of the number of bins, the upper and lower bound, and the width of the intervals, is evolving over time. The difference is that change in SEF survey is more unexpected. In the early (pre-crisis) period, the modifications in the survey design are not relevant to fluctuations of underlying variables while the adjustments in the crisis period are mostly made in response to the changing of the economic environment.

1. Inflation:

    a. From 1996Q2: [<1,1-2.5,2.5-4,4-5.5,>5.5]
    b. From 1998Q1: [<1.5,1.5-2.5,2.5-3.5,>3.5]
    c. From 1999Q2: [<1.5,1.5-2,2.5-3,3-3.5,>3.5]
    d. From 2004Q1: [<1,1-1.5,1.5-2,2-2.5,2.5-3,>3]
    e. From 2009Q1: [<0,0-1, 1-1.5,1.5-2,2-2.5,2.5-3,>3]

2. GDP growth:
    a. From 1998Q1: [<0,0-1,1-2,2-3,3-4,>4]
    b. From 2004Q1: [<1,1-2,2-3,>3]
    c. From 2008Q4: [<0,0-1,1-2,2-3,>3]
    d. From 2009Q1: [<(-1),(-1)-0,0-1,1-2,2-3,>3]

Matching

To remove the outliers in the density forecasts that are quite unreasonable and hard to justify, we continue to apply the matching procedure to refine our sample in this chapter. The matching method we use is based on the bounds implied by the density forecasts (Engelberg, Manski and Williams, 2009). The first step to construct the matched sample is to calculate the lower and upper bounds of both subjective median and mean. The interval where the median lies in can be obtained from the probabilistic responses directly. To calculate lower and upper bounds of the subjective mean, we assume that each bin's probability mass is placed at the bin's lower and upper endpoint respectively. The results are then generated by averaging the lower and upper endpoints weighted by the probabilities. If the point forecast is located within any of the two sets of bounds, the density forecast is counted as 'matched' to the point forecast. The matching takes care of another issue in the sample, that is the presence of outliers and unusual observations that appear to be errors of some sort, or at least are difficult to otherwise justify. These outliers are also removed via the matching process.

4.2.2 Z-statistics

The measure of z-statistics is constructed following the method of Probability Integral Transform (Diebold, Gunther and Tay, 1998). Suppose the true data generating process is $f_t(y_t)$ and the sequence of density forecasts is $p_t(y_t)$, the probability integral transform, z_t is defined as the cumulative density function corresponding to the density of $p_t(y_t)$ evaluated at the realization of data y_t.

$$z_t = \int_{-\infty}^{y_t} p_t(u)du = P_t(y_t)$$

If the sequence of density forecasts $\{p_t(y_t)\}_{t=1}^{m}$ coincides with the data $\{f_t(y_t|\Omega_t\}_{t=1}^{m}$, then the sequence of probability integral transforms of $\{y_t\}_{t=1}^{m}$ with respect to $\{p_t(y_t)\}_{t=1}^{m}$ is $i.i.d. U(0,1)$. That is,

$$\{z_t\}_{t=1}^{m} \overset{iid}{\sim} U(0,1)$$

In this chapter, we are not testing the i.i.d. property of z-statistic though the tests of i.i.d U(0,1) are readily available. Instead, we are focusing on the more informative outputs, the shape of distribution of z-statistic, and analyze what the different patterns of the distributions can deliver. We categorize all patterns into different groups according to two criteria: Confidence (certain/uncertain) and Bias (optimistic/pessimistic). Figure 4.1 provides an overview of the four patterns.

Confidence

When the shape of z-statistic is like a bell, it means the density forecasts may be too wide that most of z-statistics are valued at the central part of 0-100, indicating that forecasters are too uncertain and less confident in their predictions.

On the contrary, when the measures are more concentrated in the two tails, the respondents may become too confident in their expectations and thus the density forecasts have much smaller spread.

Bias

The distributions of z-statistic sometimes could be quite asymmetric with the shape being either left or right skewed. When the shape is left-skewed, there is an upward bias, indicating that forecasters may be too optimistic. On the other hand, a right-skewed distribution of z-statistic suggests that the forecasters may be too pessimistic and the forecasts are downward biased. In the next section, we will focus on the shape of z-statistic and examine how it evolves over time.

4.2.3 Current Progress

A SPF

Inflation

Confidence: Figure 4.2 depicts distributions of z-statistics for inflation density forecasts in US-SPF, with different columns illustrating different periods. The five columns from left to right indicate the following periods: 1992-2007 (full sample), 1992-2008 (before crisis), 2009 (during the crisis), 2010-2013 (after crisis and before the survey change), 2014-2017 (after the survey change). Since there is more information available at short horizon, we expect the forecasters to be more confident as target period gets closer. However, Figure 4.2 shows that it is not the case. Z-statistics are more evenly distributed in the long horizon but become concentrated in the central areas at short horizons. This indicates that though we do observe the shrink of the spread of density forecasts, the speed of shrinking is not enough compared to the information availability. In addition, forecasters are generally more uncertain after crisis. As for the survey change in

2014Q1 that is neutral to economic environmental, there is no significant change in distributions of z-statistic in short horizon. This again shows a controversial relationship between the measure of forecaster uncertainty and the confidence level revealed by the distribution of z-statistic: the reduction in forecasters uncertainty is not necessarily linked to the higher confident level or more evenly distributed z-statistic.

Bias: A sudden drop in the inflation is always linked to an upward bias in forecasts. We observe evident left-skewed shape of z-statistics in 2009 for current-year forecasts (short horizon) and in 2014 for following-year forecasts (long horizon), where there is a sudden drop in inflation in 2009 and 2015 respectively.

GDP

Figure 4.3 depicts distributions of z-statistics for output density forecasts in US-SPF. The different columns from left to right indicate different sample periods. The rows from top to bottom indicate the z-statistics relating to different horizons. During crisis, there exists upward bias in current-year forecasts (indicating the sudden depression is unexpected) and downward bias in following-year forecasts when the economy recovers, indicating the recovery process is more prompt than what forecasters expected.

After crisis, forecasters generally become more uncertain.

Almost all the biases are observed after a sudden change in the underlying economic indicators, suggesting the hysteresis in forecasters' response. This finding supports the idea of learning in expectation formations process.

B ECB-SPF

Figure 4.4 depicts distributions of z-statistics for inflation density forecasts in ECB-SPF. The different columns from left to right indicate different sample periods. The rows from top to bottom indicate the z-statistics relating to different horizons. The shapes of z-statistics distribution of ECB-SPF shows that many of the measures are valued around either 0 or 100, resulting in spikes at the two tails of the distribution. It seems to suggest the "too certain in confidence" pattern as shown in Figure 4.1, however when we looked at the heat map of z-statistics which depicts all the survey series, we find that it's not the case. Figure 4.5 shows heat map of z-statistics of all the survey series for inflation density forecasts in ECB-SPF. The rows from top to bottom indicate the z-statistics relating to different horizons. In each figure, the horizontal line indicates different survey dates and vertical line shows the corresponding range or bin of z-statistics. The darkness of each bin shows to what extent the z-statistics are clustered in it. The reason for the clustering in the two tails is that the real indicators changes so drastically that forecasters' adjustment fails to catch up with these sudden changes. This hysteresis is widely observed in the datasets, especially during the post-crisis period in Europe when the economy exhibits great volatility. Figure 4.6 and Figure 4.7 give quite similar patterns in output forecasts to those in inflation expectations.

C SEF

Not surprisingly, the hysteresis is also observed in z-statistics of SEF for both GDP growth and inflation forecasts. The sudden increase or decrease in underlying variables is always followed with z-statistics stacked in the right (100) or left (0) tails of the distributions. However, as we noted before, SEF is different in the survey design compared to US- and ECB-SPF. It changed the specification of the survey questions many times and some of the changes are neutral to the economic changes. This allows us to further examine whether forecasters' behavior can be improved by or restricted to the survey design.

Inflation

Two changes in inflation survey design are neutral to the economy:

a. At 1998Q1: from [<1,1-2.5,2.5-4,4-5.5,>5.5] to [<1.5,1.5-2.5,2.5-3.5,>3.5]

b. At 1999Q2: from [<1.5,1.5-2.5,2.5-3.5,>3.5] to [<1.5,1.5-2,2.5-3,3-3.5,>3.5]

Figure 4.8 depicts distributions of z-statistics for inflation density forecasts in SEF and Figure 4.9 shows heat map of z-statistics of all the survey series for inflation density forecasts in SEF. In both cases of survey change listed above, the z-statistic indicates better performance of density forecasts in terms of more even distribution and reasonable confidence level.

GDP

At 2004Q1, the specification for the survey question of GDP growth changed from [<0,0-1,1-2,2-3, ,3-4,>4] to [<1,1-2,2-3,>3]. This change is not relevant to the real economy, and contrariwise, the economy became even more volatile during that period. Figure 4.10 depicts distributions of z-statistics for output density forecasts in SEF and Figure 4.11 shows heat map of z-statistics of all the survey series for output density forecasts in SEF. As a result of the survey change in 2004Q1, the bounds are too tight that we observe many z-statistics stacked in the two ends of the range [0-100]. After crisis and several changes in survey design, the final setting is [-1,3] and more reasonable than [1,3], with z-statistic behaving better than pre-crisis period.

4.2.4 Future Plan

The design of the survey matters. When the interior bounds are too close to the predicted economic variables or even fail to include the underlying indicator, the density forecasts will be less informative due to the limitation of the specification of the questions. To better capture the true belief of the forecasters, a more flexible survey scheme is necessary to allow the economic experts to express precisely what is in their mind. Especially, the analysis in SEF shows that while proper refinement of the interval numbers and smaller interval width may help improve the performance of density forecasts (Inflation, 1998Q1, 1999Q2), improper shrinkage of the range or number of intervals may limit the information that density forecasts could deliver (GDP, 2004Q1).

The hysteresis in z-statistic supports the idea of learning in expectation formation process. Survey designer adjusts the question settings only after observing a sudden change in the economy. When survey respondents fail to catch up with the swift change in the macroeconomic indicators, the z-statistics would cluster in the tails of the distributions.

Lagged adjustment in periods with severe volatility also explains the controversial relationship between the measure of forecaster uncertainty and the z-statistic distributions observed in our sample. When forecaster uncertainty decreases as more inflation is available in short horizon or due to refinement in the survey design, we expect more evenly distributed z-statistics. However, the shape of z-statistics is much more complex and closely related to the variation of the real indicators.

This study provides a new application of the traditional density forecast evaluation method and adds to the literature an innovative prospective on analyzing the determinants of key features of density forecasts. My future plan is to extend the analysis to include more controls so that I can identify the effect of survey design from real economic factors on forecast performance. In addition, a deeper investigation of the link between z-statistics and other key features of density forecasts including dispersion and forecaster uncertainty may be of great help to explain the certain patterns observed in the surveys.

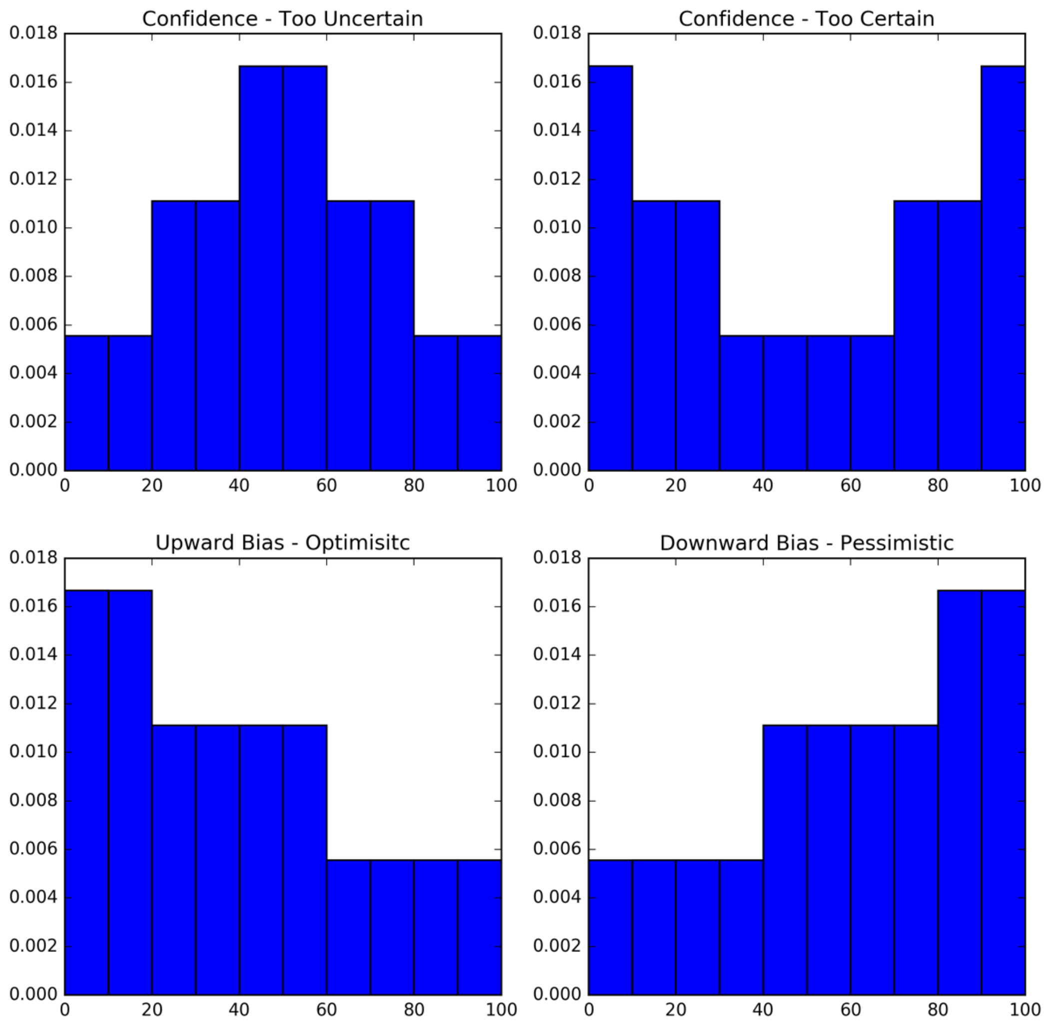

Figure 4.1 Patterns of Distributions of Z-statistic

An overview of the four patterns revealed by the distributions of Z-statistic. The top two relate to the criterion of confidence level and bottom two relate to the bias.

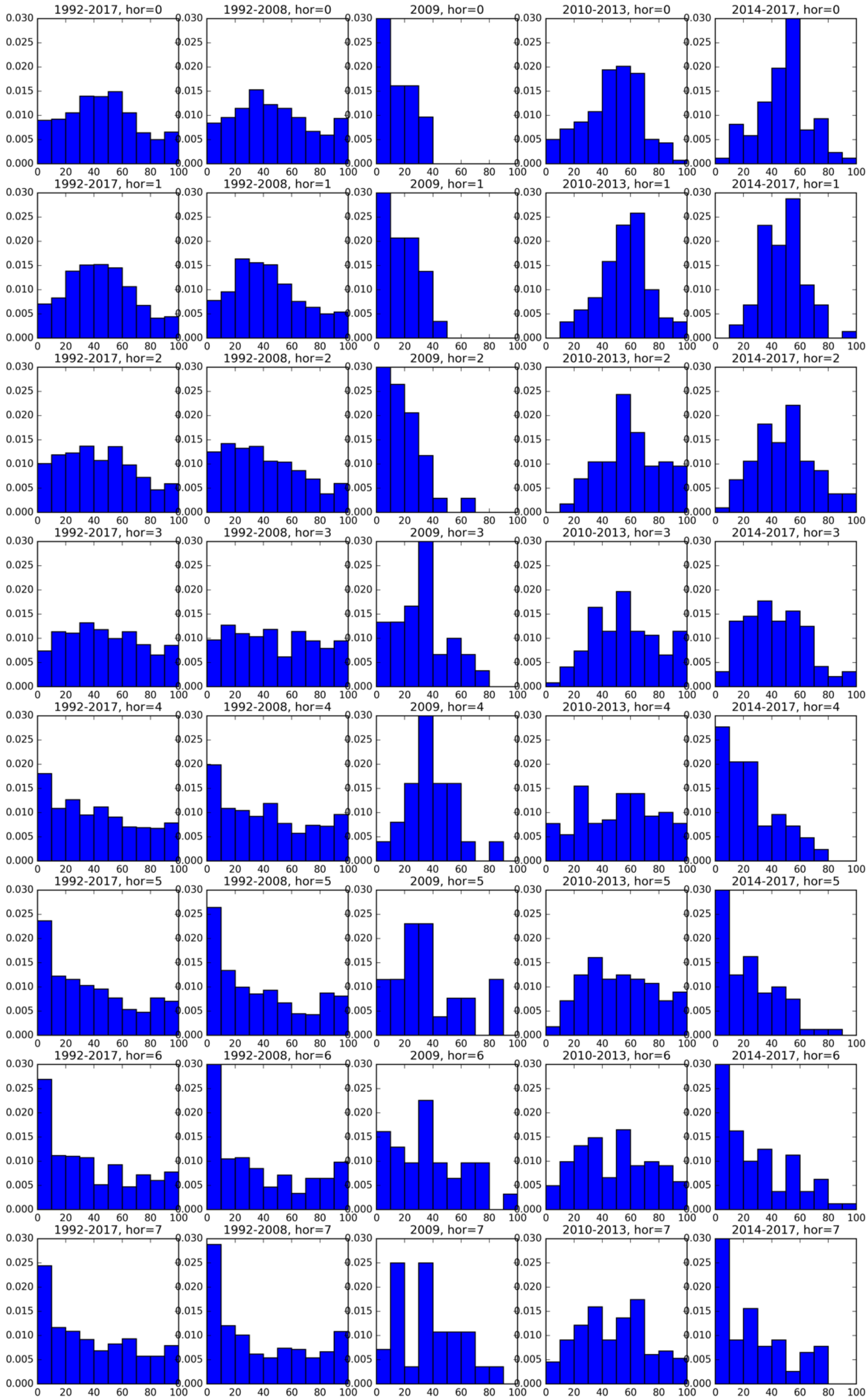

Figure 4.2 Z-statistic of US-SPF: Inflation, By Period

Distributions of z-statistics for inflation density forecasts in US-SPF, with different columns illustrating different periods. The five columns from left to right indicate the following periods: 1992-2007 (full sample), 1992-2008 (before crisis), 2009 (during the crisis), 2010-2013 (after crisis and before the survey change), 2014-2017 (after the survey change).

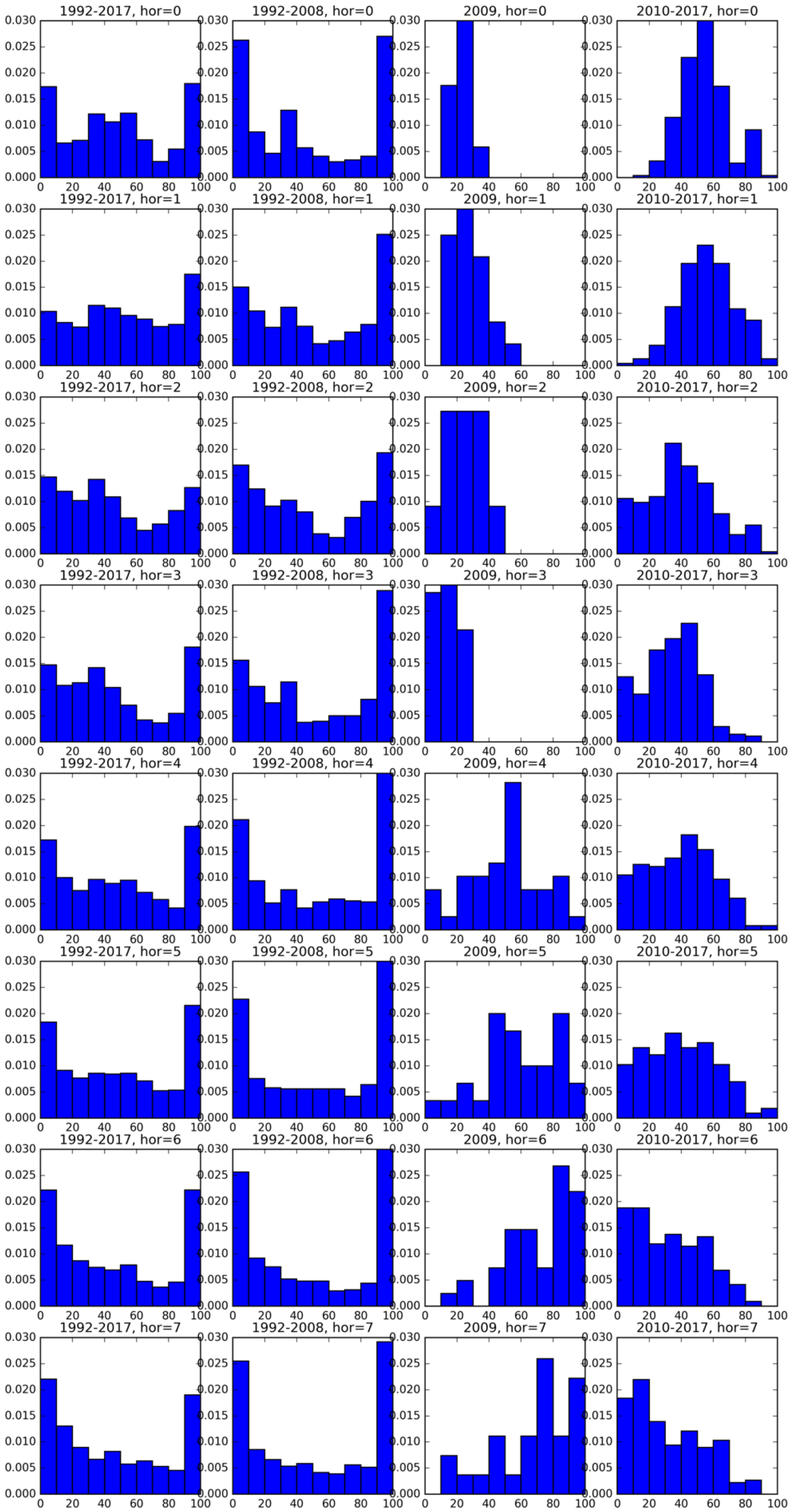

Figure 4.3 Z-statistic of US-SPF: GDP, By Period

Distributions of z-statistics for output density forecasts in US-SPF. The different columns from left to right indicate different sample periods. The rows from top to bottom indicate the z-statistics relating to different horizons.

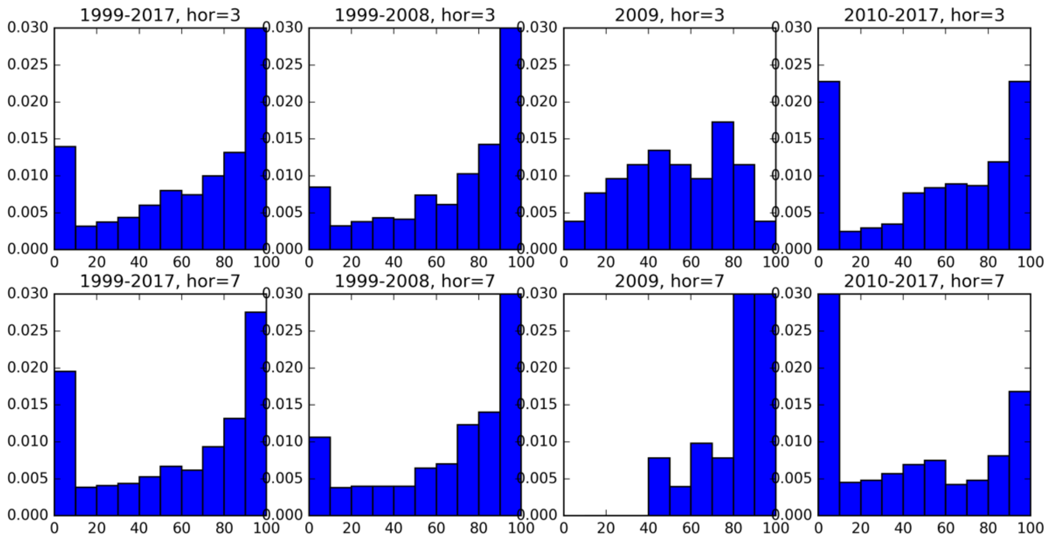

Figure 4.4 Z-statistic of ECB-SPF: Inflation, Fixed Horizon, By Period

Distributions of z-statistics for inflation density forecasts in ECB-SPF. The different columns from left to right indicate different sample periods. The rows from top to bottom indicate the z-statistics relating to different horizons.

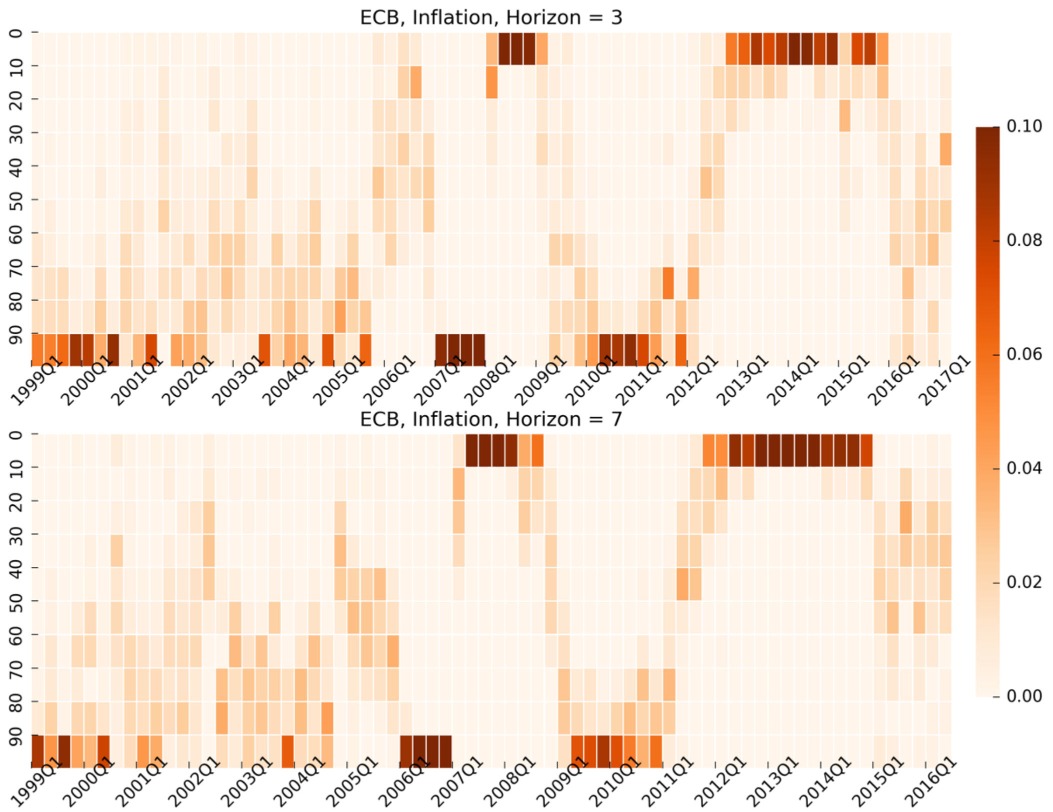

Figure 4.5 Z-statistic of ECB-SPF: Inflation, Fixed Horizon, All Survey Dates

Heat map of z-statistics of all the survey series for inflation density forecasts in ECB-SPF. The rows from top to bottom indicate the z-statistics relating to different horizons. In each figure, the horizontal line indicates different survey dates and vertical line shows the corresponding range or bin of z-statistics. The darkness of each bin shows to what extent the z-statistics are clustered in it.

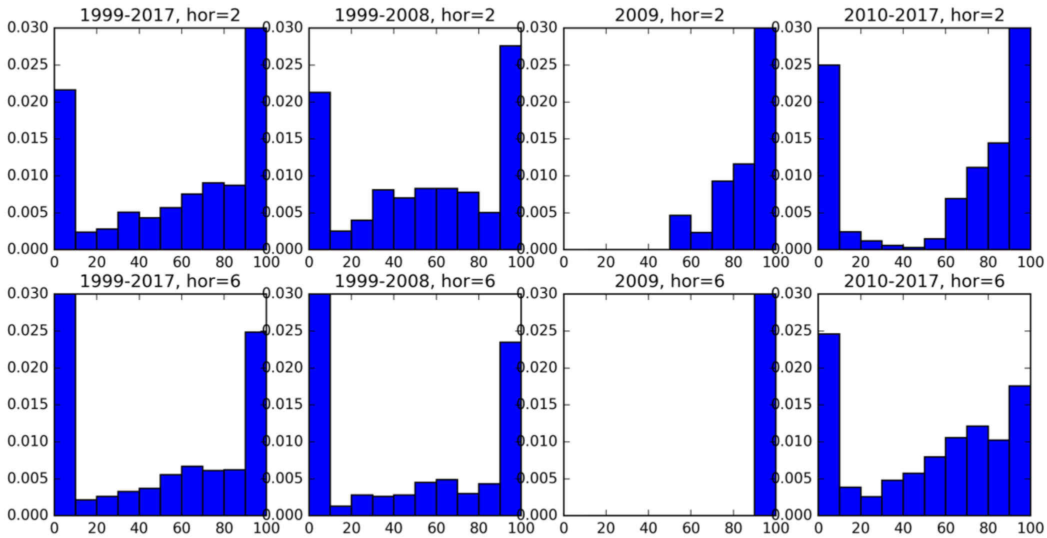

Figure 4.6 Z-statistic of ECB-SPF: GDP, Fixed Horizon, By Period

Distributions of z-statistics for output density forecasts in ECB-SPF. The different columns from left to right indicate different sample periods. The rows from top to bottom indicate the z-statistics relating to different horizons.

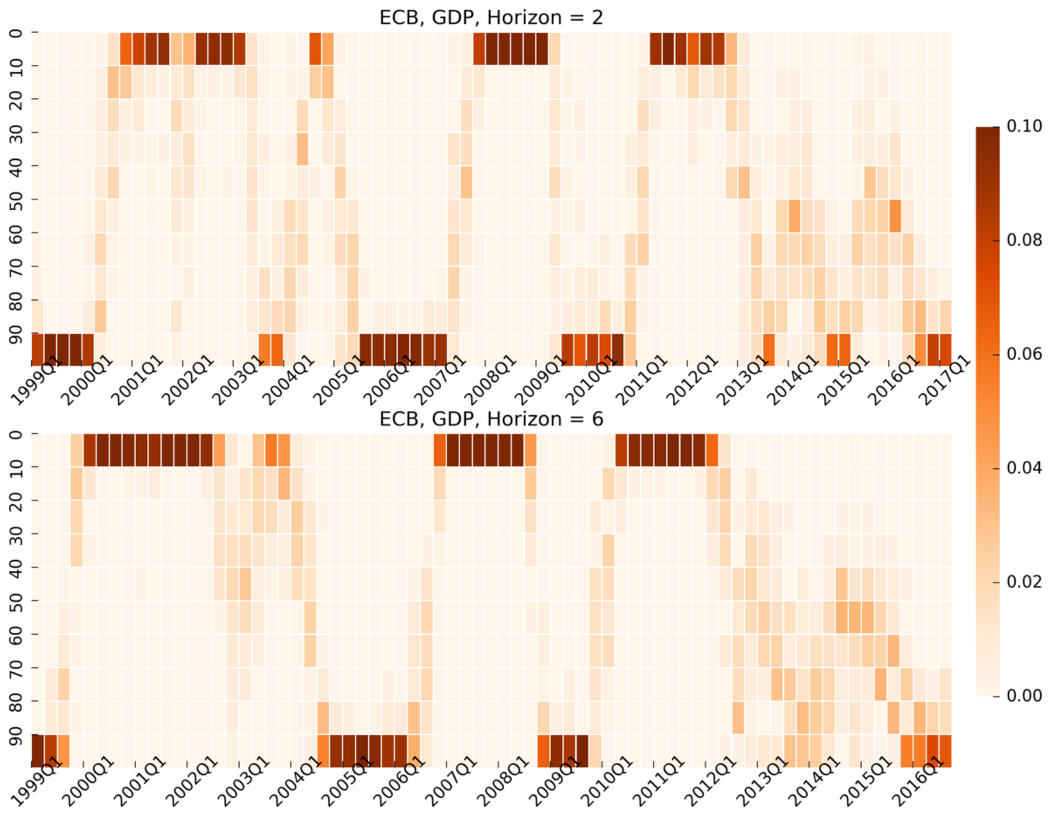

Figure 4.7 Z-statistic of ECB-SPF: GDP, Fixed Horizon, All Survey Dates

Heat map of z-statistics of all the survey series for output density forecasts in ECB-SPF. The rows from top to bottom indicate the z-statistics relating to different horizons. In each figure, the horizontal line indicates different survey dates and vertical line shows the corresponding range or bin of z-statistics. The darkness of each bin shows to what extent the z-statistics are clustered in it.

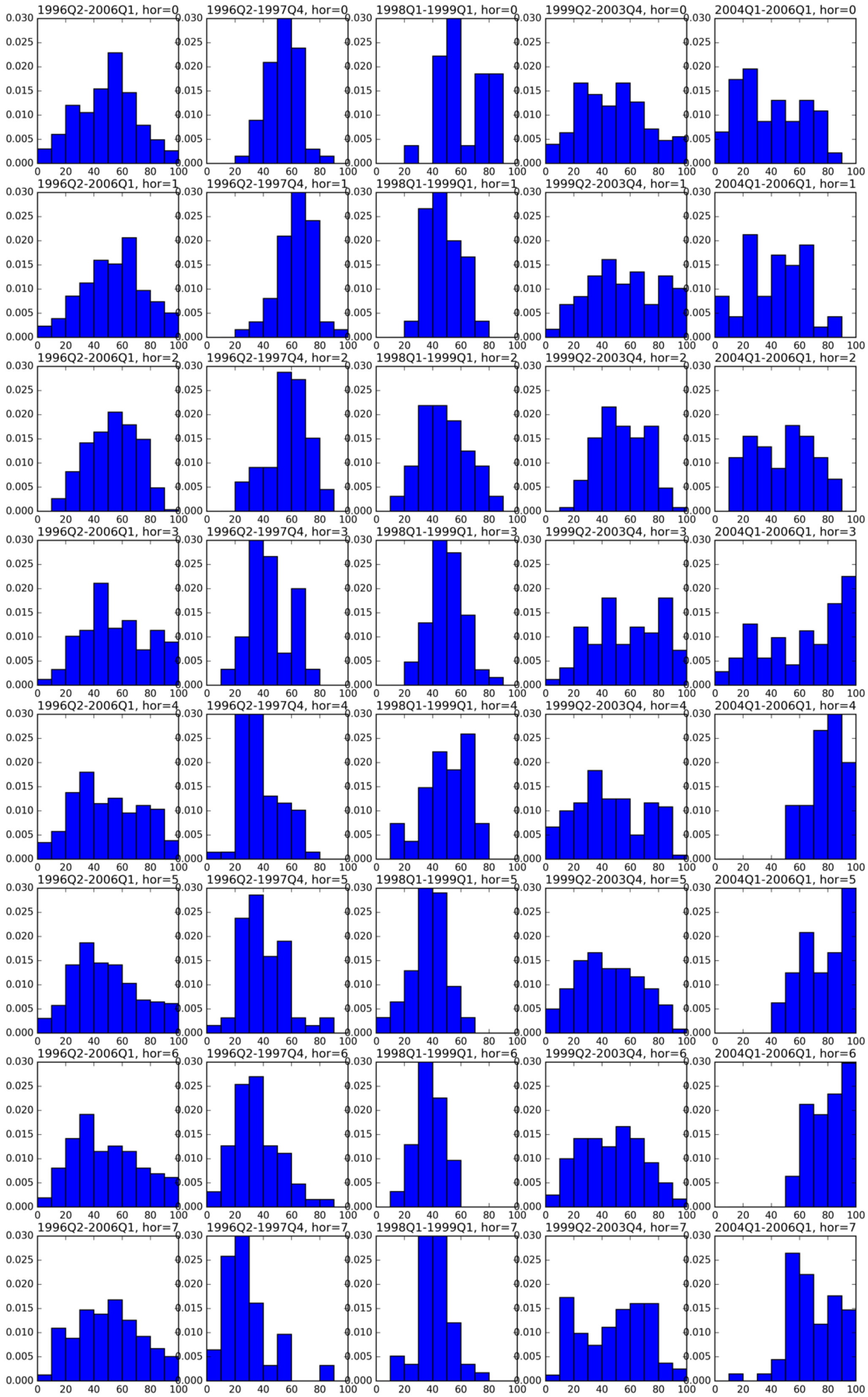

Figure 4.8 Z-statistic of SEF: Inflation, Fixed Event, By Period

Distributions of z-statistics for inflation density forecasts in SEF. The different columns from left to right indicate different sample periods. The rows from top to bottom indicate the z-statistics relating to different horizons.

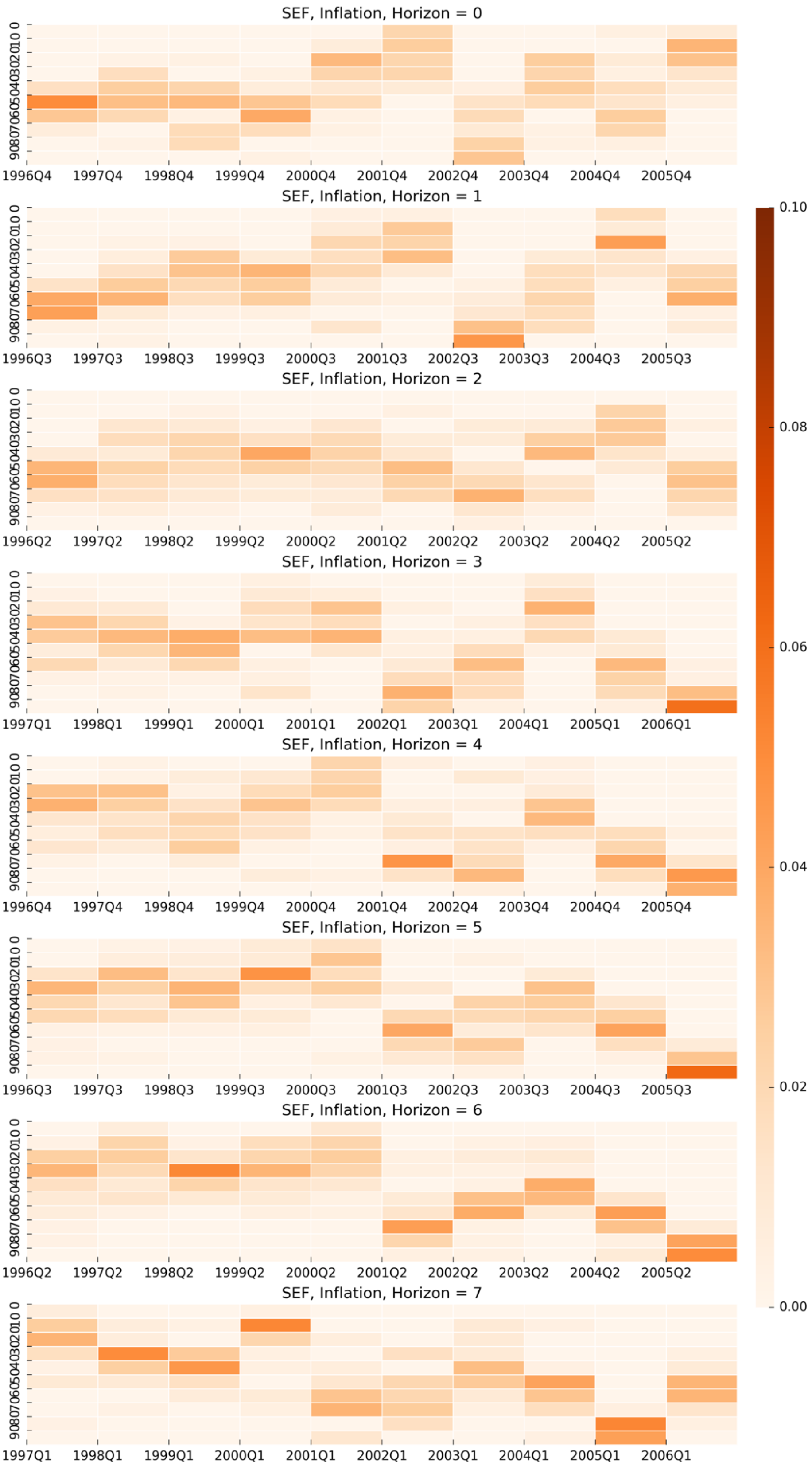

Figure 4.9 Z-statistic of SEF: Inflation, Fixed Event, All Survey Dates

Heat map of z-statistics of all the survey series for inflation density forecasts in SEF. The rows from top to bottom indicate the z-statistics relating to different horizons. In each figure, the horizontal line indicates different survey dates and vertical line shows the corresponding range or bin of z-statistics. The darkness of each bin shows to what extent the z-statistics are clustered in it.

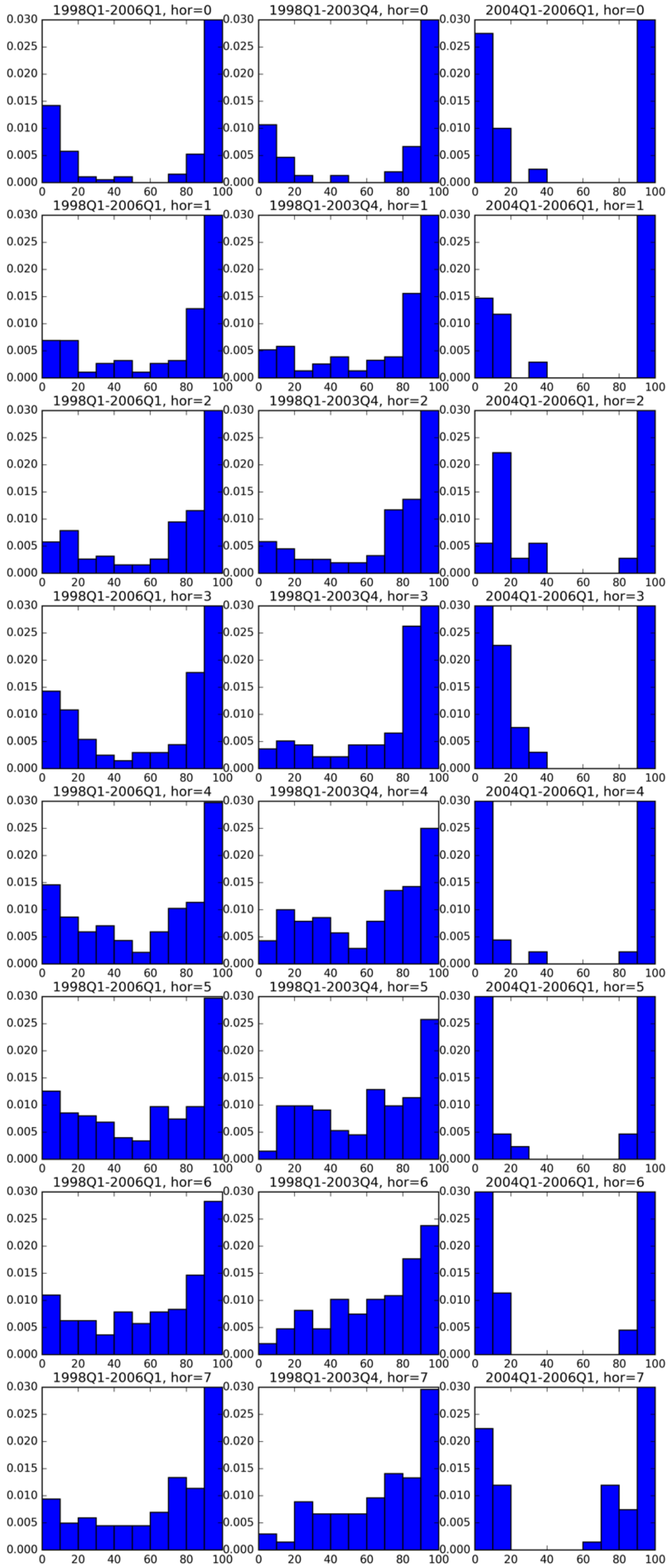

Figure 4.10 Z-statistic of SEF: GDP, Fixed Event, By Period

Distributions of z-statistics for output density forecasts in SEF. The different columns from left to right indicate different sample periods. The rows from top to bottom indicate the z-statistics relating to different horizons.

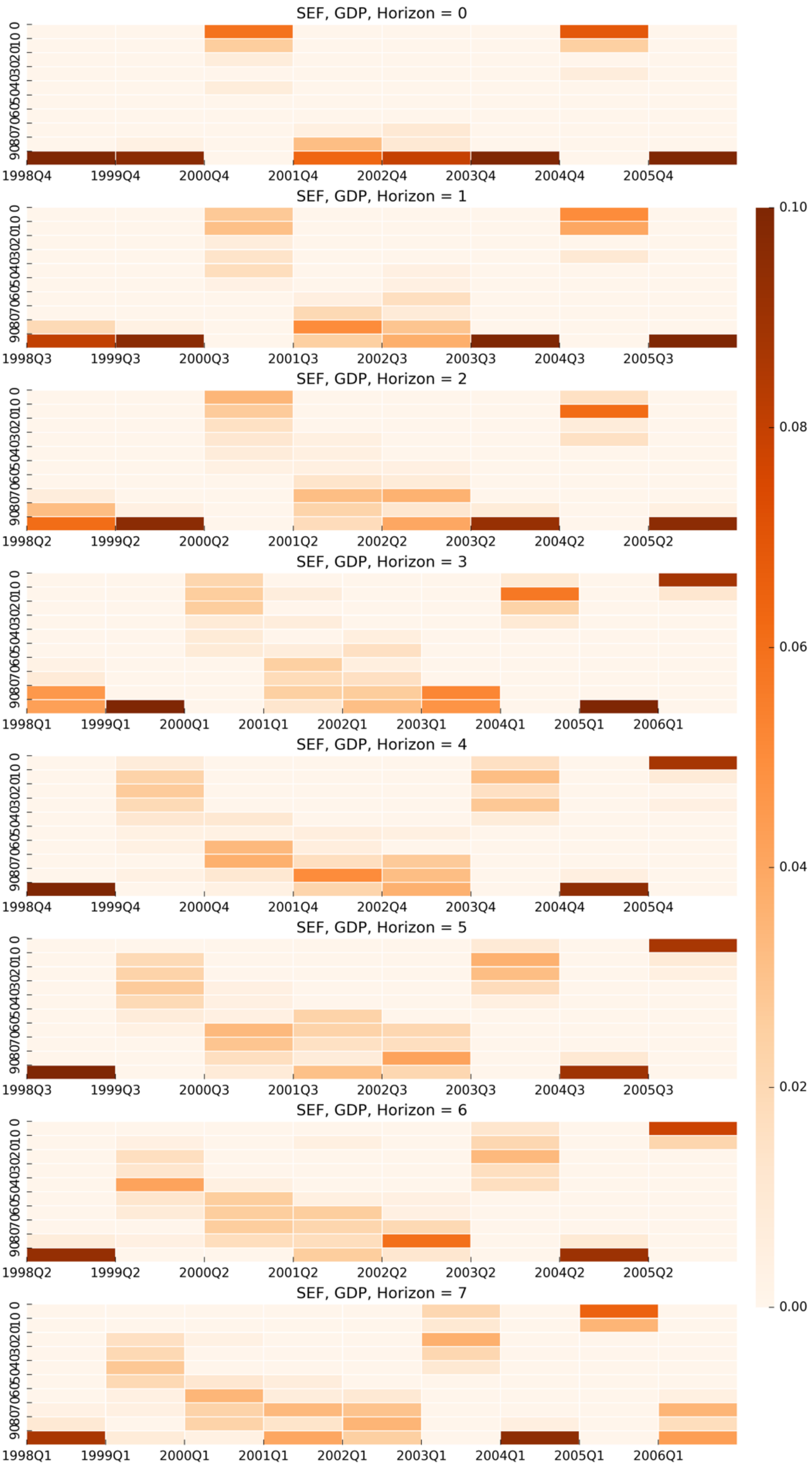

Figure 4.11 Z-statistic of SEF: GDP, Fixed Event, All Survey Dates

Heat map of z-statistics of all the survey series for output density forecasts in SEF. The rows from top to bottom indicate the z-statistics relating to different horizons.

In each figure, the horizontal line indicates different survey dates and vertical line shows the corresponding range or bin of z-statistics. The darkness of each bin shows to what extent the z-statistics are clustered in it.

References

Ackert, L. F. and G. Athanassakos (1997), "Prior Uncertainty, Analyst Bias, and Subsequent Abnormal Returns", Journal of Financial Research, 20(2), 263-273.

Ang, A., G. Bekaert, and M. Wei (2007), "Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?", Journal of Monetary Economics, 54(4), 1163-1212.

Ang, A., M. Piazzesi, and M. Wei (2006), "What Does the Yield Curve Tell Us About GDP Growth?", Journal of Econometrics, 131(1-2), 359–403.

Baker, S.R., N. Bloom, and S.J. Davis (2016), "Measuring Economic Policy Uncertainty", The Quarterly Journal of Economics, 131(4), 1593-1636.

Ball, L. (1992), "Why Does Higher Inflation Raise Inflation Uncertainty?", Journal of Monetary Economics, 29(3), 371-388.

Boero, G., J. Smith, and K.F. Wallis (2008), "Uncertainty and Disagreement in Economic Prediction: The Bank of England Survey of External Forecasters", The Economic Journal, 118(530), 1107-1127.

Boero, G., J. Smith, and K.F. Wallis (2015), "The Measurement and Characteristics of Professional Forecasters' Uncertainty", Journal of Applied Econometrics, 30(7), 1029-1046.

Bourveau, T. and K. Law (2016), "Katrina and Analysts", SSRN Working Paper

Capistran, C. and A. Timmermann (2009), "Disagreement and Biases in Inflation Expectations", Journal of Money, Credit and Banking, 41(2-3), 365-396.

Clark, A., A. Ghosh and S. Hanes (2018), "Inflation Expectations in Singapore: a Behavioural Approach", Macroeconomic Review (April 2018), 89-98.

Das, S., C. B. Levine and K. Sivaramakrishnan (1998), “Earnings Predictability and Bias in Analysts' Earnings Forecasts”, Accounting Review, 277-294.

Diebold, F. X., Gunther. T. A., and Tay, A. S. (1998), “Evaluating Density Forecasts with Applications to Financial Risk Managment”, International Economic Review, 39(4), 863-883.

Dolvin, S. D., M. K. Pyles and Q. Wu (2009), “Analysts get SAD too: The effect of seasonal affective disorder on stock analysts' earnings estimates”, The Journal of Behavioral Finance, 10(4), 214-225.

Engelberg, J., C.F. Manski, and J. Williams (2009), “Comparing the Point Predictions and Subjective Probability Distributions of Professional Forecasters”, Journal of Business & Economic Statistics, 27(1), 30-41.

Estrella, A. and G.A. Hardouvelis (1991), “The Term Structure as Predictor of Real Economic Activity”, The Journal of Finance, 46(2), 555-576.

Estrella, A. and F.S. Mishkin (1998), “Predicting U.S. Recessions: Financial Variables as Leading Indicators”, The Review of Economics and Statistics, 80(1), 45-61.

Giordani, P. and P. Soderlind (2003), “Inflation Forecast Uncertainty”, European Economic Review, 47(6), 1037-1059.

Hamilton, J.D. and D.H. Kim (2002), “A Re-Examination of the Predictability of Economic Activity Using the Yield Spread”, Journal of Money, Credit and Banking, 34(2), 340-360.

Jurado, K., S.C. Ludvigson, and S. Ng (2015), “Measuring Uncertainty”, The American Economic Review, 105(3), 1177-1216.

Kandel, E. and B-Z. Zilberfarb (1999), “Differential Interpretation of Information in Inflation Forecasts”, The Review of Economics and Statistics, 81(2), 217-226.

Lahiri, K. and F. Liu (2006), “Modelling Multi-Period Inflation Uncertainty Using a Panel of Density Forecasts”, Journal of Applied Econometrics, 21(8), 1199-1219.

Lo, K. and S. S. Wu (2017). “The Impact of Seasonal Affective Disorder on Financial Analysts”, The Accounting Review.

Ludvigson, Sydney C., Sai Ma, and Serena Ng (2015). “Uncertainty and business cycles: exogenous impulse or endogenous response?”, No. w21803, National Bureau of Economic Research.

Mankiw, N.G., R. Reis and J. Wolfers (2003), “Disagreement about Inflation Expectations”, NBER Macroeconomic Annual, 18, 209-248.

Manzan, S. (2011), “Differential Interpretation in the Survey of Professional Forecasters”, Journal of Money, Credit and Banking, 43(5), 993-1017.

Patton, A.J. and A. Timmermann (2007), “Testing Forecast Optimality under Unknown Loss”, Journal of the American Statistical Association, 102(480), 1172-1184.

Patton, A.J. and A. Timmermann (2010), “Why Do Forecasters Disagree? Lessons from the Term Structure of Cross-Sectional Dispersion”, Journal of Monetary Economics, 57(7), 803-820.

Rich, R. and J. Tracy (2010), “The Relationships among Expected Inflation, Disagreement, and Uncertainty: Evidence from Matched Point and Density Forecasts”, The Review of Economics and Statistics, 92(1), 200-207.

Rossi, B. and T. Sekhposyan (2010), "Have Economic Models' Forecasting Performance for US Output Growth and Inflation Changed over Time, and When?", International Journal of Forecasting, 26(4), 808-835.

Stock, J. H. and M.W. Watson (2009), "Phillips Curve Inflation Forecasts", in J. Fuhrer, Y. Kodrzycki, J. Little, G. Olivei (eds), Understanding Inflation and the Implications for Monetary Policy, 99-202, Cambridge, MIT Press.

Zarnowitz, V. and L.A. Lambros (1987), "Consensus and Uncertainty in Economic Prediction", Journal of Political Economy, 95(3), 591-621.

Zhang, X. F. (2006), "Information uncertainty and analyst forecast behavior", Contemporary Accounting Research, 23(2), 565-590.